



XX Claim 7, Page 21; 44pp; English.  
 PS  
 CC The present invention relates to dental products such as toothpastes,  
 CC mouthwash and dental floss comprising a base material and a compound  
 CC which promotes bone growth. Such compounds are peptide sequences  
 CC comprising 10-50 amino acids and containing an integrin binding  
 CC motif such as RGD in the D- or L- form, preferably the L-configuration.  
 CC The peptides of the invention are useful for treating or preventing  
 CC skeletal diseases such as dental disease. The peptides enhance tooth  
 CC and/or alveolar bone growth on areas where deterioration has occurred,  
 CC as well as the growth or recruiting of osteoblast or odontoblast cells  
 CC on the surface of the new skeletal growth. AAU93681-AAU93726 represent  
 CC bone growth enhancing peptides for use in dental products.  
 CC  
 SQ Sequence 23 AA:

Query Match 100.0%; Score 126; DB 23; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-12;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDQERGDNDISFSGDGPFKD 23  
 ID 1 TDQERGDNDISFSGDGPFKD 23  
 DB 1 TDQERGDNDISFSGDGPFKD 23

RESULT 2  
 AAO20379  
 ID AAO20379 standard: peptide: 23 AA.

AC AAO20379;  
 XX  
 DT 31-MAY-2002 (first entry)

DE C-terminal amidated synthesised peptide D-00006.

XX Bone growth; RGD motif; integrin binding motif; calcium binding motif;  
 KM glycosaminoglycan binding motif; bone loss; renal phosphate excretion;  
 KM alveolar teeth; odontoblast; osteoclast; dental tissue; skeletal loss;  
 KW weakness; D00006.

XX Synthetic.

XX Key Location/Qualifiers  
 FT Modified-site 23  
 FT /note- "C-terminal amide"

PN WO200214360-A1.

XX 21-FEB-2002.

PD 14-AUG-2001; 2001WO-US25542.

PF 16-AUG-2000; 2000US-0641034.

PR 19-MAR-2001; 2001US-0812485.

XX (BIGB-) BIG BEAR BIO INC.

XX Kumagai Y, Blacher RW, Yoneda T;

XX WPI: 2002-291971/33.

XX New peptide compound useful for reducing bone loss, is capable of  
 PT enhancing bone growth, and comprises an integrin binding motif,  
 PT glycosaminoglycan binding motif or a calcium binding motif  
 XX  
 XX Example 1; Page 15; 50pp; English.

XX The invention relates to a peptide compound capable of enhancing bone  
 CC growth, and comprising 10-50 amino acids in a sequence, where the amino  
 CC acids are in D- or L- conformation and the sequence comprises a motif  
 CC selected from an integrin binding motif, a glycosaminoglycan binding  
 CC motif and a calcium binding motif. The peptide of the invention is useful

CC for reducing bone loss and for reducing renal phosphate excretion in an  
 CC individual. The peptide is useful for promoting regeneration of alveolar  
 CC bone and/or teeth, and increases the number and activity of odontoblasts  
 CC /osteoclasts that help form dental tissues. The peptide is also useful  
 CC for treating or preventing a condition associated with skeletal loss or  
 CC weakness. This sequence represents a C-terminal amidated synthesised  
 CC peptide D-00006 of the invention.  
 CC  
 SQ Sequence 23 AA:

Query Match 100.0%; Score 126; DB 23; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-12;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDQERGDNDISFSGDGPFKD 23  
 ID 1 TDQERGDNDISFSGDGPFKD 23  
 DB 1 TDQERGDNDISFSGDGPFKD 23

RESULT 3  
 AAU93681  
 ID AAU93681 standard: protein: 97 AA.

XX AAU93681;

DT 02-JUL-2002 (first entry)

DE Dental product bone growth enhancing peptide #1.

XX Dental product; toothpaste; mouthwash; dental floss; bone growth;  
 KM integrin binding motif; RGD; skeletal disease; dental disease; tooth;  
 KM alveolar bone growth; osteoblast; odontoblast; osteopathic.

XX Synthetic.

PN WO200213775-A1.

XX 21-FEB-2002.

PF 09-AUG-2001; 2001WO-US25101.

PR 16-AUG-2000; 2000US-225879P.

XX (BIGB-) BIG BEAR BIO INC.

XX Yoneda T, Nomizu M, Kumagai Y;

XX WPI: 2002-329525/36.

XX Dental product useful for treating skeletal diseases e.g. dental  
 PT diseases comprises a base material and a compound comprising specific  
 PT amino acid in a sequence containing the integrin binding motif  
 XX  
 XX Disclosure; Page 11; 44pp; English.

XX The present invention relates to dental products such as toothpastes,  
 CC mouthwash and dental floss comprising a base material and a compound  
 CC which promotes bone growth. Such compounds are peptide sequences  
 CC comprising 10-50 amino acids and containing an integrin binding  
 CC motif such as RGD in the D- or L- form, preferably the L-configuration.  
 CC The peptides of the invention are useful for treating or preventing  
 CC skeletal diseases such as dental disease. The peptides enhance tooth  
 CC and/or alveolar bone growth on areas where deterioration has occurred,  
 CC as well as the growth or recruiting of osteoblast or odontoblast cells  
 CC on the surface of the new skeletal growth. AAU93681-AAU93726 represent  
 CC bone growth enhancing peptides for use in dental products.  
 CC  
 SQ Sequence 97 AA:

Query Match 100.0%; Score 126; DB 23; Length 97;  
 Best Local Similarity 100.0%; Pred. No. 9.1e-12;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDQERGNDISPFGDGPFD 23  
 DB 43 TDQERGNDISPFGDGPFD 65

RESULT 4  
 ID AAO20331 standard; protein: 97 AA.

AC AAO20331;

DT 31-MAY-2002 (first entry)

DE Protein of matrix extracellular phosphoglycoprotein containing RGD #1.  
 KW Bone growth; RGD motif; integrin binding motif; calcium binding motif;  
 KW glycosaminoglycan binding motif; bone loss; renal phosphate excretion;  
 KW alveolar; teeth; odontoblast; osteoclast; dental tissue; skeletal loss;  
 KW weakness; matrix extracellular phosphoglycoprotein.

OS Unidentified.

PN MO20214360-A1.

PD 21-FEB-2002.

PE 14-AUG-2001; 2001WO-US25542.

PR 16-AUG-2000; 2000US-0641034.

PR 19-MAR-2001; 2001US-0812485.

PA (BIGB-) BIG BEAR BIO INC.

PI Kumagai Y, Blacher RW, Yoneda T;

DR WPI; 2002-291971/33.

PT New peptide compound useful for reducing bone loss, is capable of  
 enhancing bone growth, and comprises an integrin binding motif,  
 PT glycosaminoglycan binding motif or a calcium binding motif -

PS Disclosure; Page 11; 50pp; English.

CC The invention relates to a peptide compound capable of enhancing bone  
 CC growth, and comprising 10-50 amino acids in a sequence, where the amino  
 CC acids are in D- or L- conformation and the sequence comprises a motif  
 CC selected from an integrin binding motif, a glycosaminoglycan binding  
 CC motif and a calcium binding motif. The peptide of the invention is useful  
 CC for reducing bone loss and for reducing renal phosphate excretion in an  
 CC individual. The peptide is useful for promoting regeneration of alveolar  
 CC bone and/or teeth, and increases the number and activity of odontoblasts  
 CC /osteoclasts that help form dental tissues. The peptide is also useful  
 CC for treating or preventing a condition associated with skeletal loss or  
 CC weakness. This sequence represents a protein of a matrix extracellular  
 CC phosphoglycoprotein containing an RGD motif of the invention.

SO Sequence 97 AA;

Query Match 100.0%; Score 126; DB 23; Length 97;  
 Best Local Similarity 100.0%; Pred. No. 9; 1e-12;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDQERGNDISPFGDGPFD 23  
 DB 43 TDQERGNDISPFGDGPFD 65

RESULT 5  
 ID AAY53812 standard; Protein: 430 AA.  
 AC AAY53812;

DT 22-FEB-2000 (first entry)  
 DE Amino acid sequence of a human phosphatonin polypeptide.

KW Human; phosphatonin; Metastatic-tumour Excreted Phosphaturic-Element;  
 KW MEPE; Na+-dependent phosphate cotransport; vitamin D metabolism;  
 KW bone mineralisation; phosphate metabolism related disease;  
 KW hyperphosphatemia; renal osteodystrophy; renal dialysis;  
 KW secondary hyperparathyroidism; osteitis fibrosa cystica; hypercalcaemia;  
 KW x-linked hypophosphatemic rickets; hereditary hypophosphatemic rickets;  
 KW hypomineralised bone lesion; stunted growth; cystic fibrosis;  
 KW oncogenic hypophosphatemics osteomalacia; renal phosphate leakage;  
 KW renal osteodystrophy; osteoporosis; vitamin D resistant rickets;  
 KW end organ resistance; renal Fanconi syndrome; autosomal rickets;  
 KW Paget's disease; kidney failure; renal tubular acidosis; spure.

OS Homo sapiens.

PN Key

FT Modified-site  
 FT Location/Qualifiers  
 FT 8..10  
 FT /note- "protein kinase C phosphorylation site"

FT 8..11  
 FT /note- "Caseln kinase II phosphorylation site"

FT 16..21  
 FT /note- "myristoylation site"

FT 40..47  
 FT /note- "Tyrosine kinase phosphorylation site"

FT 77..79  
 FT /note- "protein kinase C phosphorylation site"

FT 118..120  
 FT /note- "protein kinase C phosphorylation site"

FT 119..224  
 FT /note- "myristoylation site"

FT 139..142  
 FT /note- "Caseln kinase II phosphorylation site"

FT 143..148  
 FT /note- "myristoylation site"

FT 152..154  
 FT /note- "cell attachment tripeptide"

FT 161..165  
 FT /note- "glycosaminoglycan attachment site"

FT 177..180  
 FT /note- "Caseln kinase II phosphorylation site"

FT 194..197  
 FT /note- "Caseln kinase II phosphorylation site"

FT 199..202  
 FT /note- "Caseln kinase II phosphorylation site"

FT 203..205  
 FT /note- "protein kinase C phosphorylation site"

FT 224..227  
 FT /note- "Caseln kinase II phosphorylation site"

FT 228..231  
 FT /note- "Caseln kinase II phosphorylation site"

FT 228..230  
 FT /note- "protein kinase C phosphorylation site"

FT 238..241  
 FT /note- "Caseln kinase II phosphorylation site"

FT 266..271  
 FT /note- "myristoylation site"

FT 291..296  
 FT /note- "myristoylation site"

FT 311..313  
 FT /note- "myristoylation site"

FT 312..314  
 FT /note- "protein kinase C phosphorylation site"

FT 315..320  
 FT /note- "protein kinase C phosphorylation site"

FT 319..321  
 FT /note- "myristoylation site"

FT 325..328  
 FT /note- "protein kinase C phosphorylation site"

FT 370..373  
 FT /note- "amidation site"

FT	Modified-site	382..386	/note= "Asu-glycosylation site"
FT	Modified-site	383..387	/note= "Asu-glycosylation site"
FT	Modified-site	384..386	/note= "protein kinase C phosphorylation site"
FT	Modified-site	389..394	/note= "myristoylation site"
FT	Modified-site	403..405	/note= "protein kinase C phosphorylation site"
FT	Modified-site	405..408	/note= "cAMP and cGMP dependent protien kinases phosphorylation site"
FT	Modified-site	408..410	/note= "protein kinase C phosphorylation site"
FT	Modified-site	409..411	/note= "protein kinase C phosphorylation site"
FT	Modified-site	423..426	/note= "protein kinase C phosphorylation site"
FT	Modified-site	425..428	/note= "Casein kinase II phosphorylation site"
FT	Modified-site	427..430	/note= "Casein kinase II phosphorylation site"
FT	Modified-site		/note= "Casein kinase II phosphorylation site"

PN	MO9960017-A2.	
XX		
PD	25-NOV-1999	99MO-EF03403.
PF	18-MAY-1999;	99MO-EF03403.
XX		
PR	18-MAY-1998;	98GB-0010681.
XX		
PR	04-SEP-1998;	98GB-0019387.
XX		
PA	(UNIO) UNIV COLLEGE LONDON.	

AA  
PI      Rowe P;  
XX  
DR -      WPI: 2000-053262/04.  
DR      N-PSDB; AAZ36447.

PT New polypeptides involved in the regulation of phosphate metabolism  
PT useful for diagnosing and treating disorders related to phosphate metabolism

PS Claim 6; Fig 8; 136pp; English.

The present sequence represents a phosphatonin polypeptide (also called Metastatic-tumour Excreted Phosphaturic-Element (MEPE)). The level of phosphatonin in a subject modulates Na<sup>+</sup>-dependent phosphate cotransport vitamin D metabolism and/or bone mineralisation. The phosphatonin polypeptides, polynucleotides, vectors and antibodies are used to treat phosphate metabolism related disease. They are used for treatment of hyperphosphatemia, or renal osteodystrophy, hyperphosphatemia in renal dialysis/pre-dialysis, secondary hyperparathyroidism or osteitis fibrosa cystica, or x-linked hypophosphatemic rickets, hereditary hypophosphatemic rickets with hypocalcaemia (HNHR), hypomineralised bone lesions, stunted growth in juveniles, oncogenic hypophosphatemic osteomalacia, renal phosphate leakage, renal osteodystrophy, osteoporosis, vitamin D resistant rickets, end organ resistance, renal Fanconi syndrome, autosomal rickets, Paget's disease, kidney failure, renal tubular acidosis, cystic fibrosis or sprue. The polypeptide may also be used to manufacture combined preparations for simultaneous, separate or sequential use for the treatment of phosphate metabolism disorders. A transformed osteoblast or bone cell line capable of phosphatonin overexpression is useful for the production of phosphatonin.

**SQ Sequence 430 AA;**

Query Match	100.0%	Score 126:	DB 21:	Length 430:
Best Local Similarity	100.0%	Pred. No. 5e-11:		
Matches 23; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

```
Qy      1  TDLQERGDNDISPFSGDGQPFKD  23
          |||||
Db     147  TDLQERGDNDISPFSGDGQPFKD  169
```

RESULT 6  
AAB62669  
ID AAB62669 standard; Protein; 430 AA

AC AAB62669

DT 23-JUL-2001 (first entry)

Truncated phosphatoin polypeptide (truncated MEPE).

KM Metastatic-tumour excreted phosphatonic element; MEPE, phosphatonin;  
KM phosphate; vitamin-D; skeletal formation; mineralization; truncated;  
KM osteopathic; antigen; cytosolic; human.

OS Homo sapiens.

FH	Key	Location/Qualifiers
FM	W43-44660000	161 103

FT /note- "the amino acid residues in this region are  
FT indicated incorrectly in the sequence provided  
FT in the sequence listing"

PN WO200132878-A2.

PD 10-MAY-2001.

PF 31-OCT-2000; 2000WO-EP10747.

PR	04-NOV-1999;	99US-0434185
PR	08-NOV-1999;	99CD-0035434

XX  
PA  
(UNITED) COLLEGE LONDON

PI ROWE P.

DR WPI; 2001-343487/36.

DR N-PSDB; AAF83759.

PT New phosphatonin polypeptide a regulator of phosphate metabolism, for  
PT diagnosing and treating disorders of phosphate, vitamin-D metabolism  
PT skeletal formation e.g. osteoporosis, Paget's disease, gout -

PS Claim 4; Fig 8; 135pp; English.

The invention relates to a novel human protein, metastatic-tumour excreted phosphaturic element (MTEP) or phosphatonin (modulator of CC phosphate and vitamin-D metabolism). The phosphatonin (modulator of CC polynucleotides and specific antihobdes are useful for treating a CC disorder of phosphate or vitamin D metabolism, skeletal formation and mineralization. Phosphatonins are used to treat hyperphosphatemia, renal osteodystrophy, secondary hyperparathyroidism, osteitis fibrosa cystica or gout. It is used to prepare a medicament for treating X-linked CC hypophosphatemic rickets, hereditary hypophosphatemic rickets with hypocalcemia (HHRH), hypomineralized bone lesions, stunted growth in CC juveniles, oncogenic hypophosphatemic osteomalacia, renal phosphate leakage, renal osteodystrophy, osteoporosis, vitamin-D resistant rickets, CC end organ resistance, renal Fanconi syndrome, autosomal recessive, Paget's disease, kidney failure, renal tubular acidosis, cystic fibrosis or CC sprue. Phosphatonin polynucleotides are useful as molecular weight markers on Southern gels, as diagnostic probes for detecting the presence CC of a specific mRNA. Phosphatonin polypeptides are also useful for CC identifying agonists and antagonists, compounds which bind to CC phosphatonin and drug candidates for therapy of phosphate metabolism CC disorders. The present sequence represents a truncated form of CC phosphatonin (MTEP).

**SQ Sequence 430 AA;**



QY 1 TDLQERGDNDISPFSGDGPFFKD 23  
 DB 242 TDLQERGDNDISPFSGDGPFFKD 264

RESULT 9  
 AAB62689 standard; Protein: 525 AA.

AC AAB62689;  
 XX  
 DT 23-JUL-2001 (first entry)

XX Phosphatonin polypeptide (MEPE).  
 DE Metastatic tumour excreted phosphaturic element; MEPE; phosphatonin;  
 XX phosphatase; vitamin-D; skeletal formation; mineralization; truncated;  
 KM osteopachic; antlygout; cytostatic; human.  
 XX Homo sapiens.

OS MO200132878-A2.

PN 10-MAY-2001.

PD 31-OCT-2000; 2000MO-EP10747.

PF 04-NOV-1999; 99US-0434185.  
 PR 08-NOV-1999; 99GB-0026424.

PA (UNLO ) UNIV COLLEGE LONDON.

PI Rowe P;

DR WPI: 2001-343487/36.  
 XX N-PSDB: AAF83764.

PT New phosphatonin polypeptide a regulator of phosphate metabolism, for  
 PT diagnosing and treating disorders of phosphate, vitamin-D metabolism,  
 PT skeletal formation e.g. osteoporosis, Paget's disease, gout

PS Claim 4: Page 133-134; 135pp; English.

XX The invention relates to a novel human protein, metastatic-tumour  
 CC excreted phosphaturic element (MEPE) or phosphatonin (modulator of  
 CC phosphate and vitamin-D metabolism). The phosphatonin polypeptides,  
 CC polynucleotides and specific antibodies are useful for treating a  
 CC disorder of phosphate or vitamin D metabolism, skeletal formation and  
 CC mineralization. Phosphatonins are used to treat hyperphosphatemia, renal  
 CC osteodystrophy, secondary hyperparathyroidism, osteitis fibrosa cystica  
 CC or gout. It is used to prepare a medicament for treating x-linked  
 CC hypophosphatemic rickets, hereditary hypophosphatemic rickets with  
 CC hypercalcaemia (HHRH), hypomineralized bone lesions, stunted growth in  
 CC juveniles, oncogenic hypophosphatemic osteomalacia, renal phosphate  
 CC leakage, renal osteodystrophy, osteoporosis, vitamin-D resistant rickets,  
 CC end organ resistance, renal Fanconi syndrome, autosomal rickets, Paget's  
 CC disease, kidney failure, renal tubular acidosis, cystic fibrosis or  
 CC sprue. Phosphatonin polynucleotides are useful as molecular weight  
 CC markers on Southern gels, as diagnostic probes for detecting the presence  
 CC of a specific mRNA. Phosphatonin polypeptides are also useful for  
 CC identifying agonists and antagonists, compounds which bind to  
 CC phosphatonin and drug candidates for therapy of phosphate metabolism  
 CC disorders. The present sequence represents the amino acid sequence of  
 CC the entire phosphatonin (MEPE).

XX Sequence 525 AA;

Query Match 100.0%; Score 126; DB 22; Length 525;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-11;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDLQERGDNDISPFSGDGPFFKD 23  
 DB 242 TDLQERGDNDISPFSGDGPFFKD 264

DB 242 TDLQERGDNDISPFSGDGPFFKD 264

RESULT 10  
 ABB08526 standard; Protein: 525 AA.

AC ABB08526;  
 XX  
 DT 23-MAY-2002 (first entry)

XX Human phosphatonin.

DE Human phosphatonin; cytostatic; antidiabetic; antiinflammatory;  
 KM hyperphosphatemia; arteriosclerosis; heart failure;  
 KM diabetic renal disease; kidney failure; cystic fibrosis.

OS Homo sapiens.

PN MO200198495-A1.

PD 27-DEC-2001.

PF 20-JUN-2001; 2001MO-JP05263.

PR 21-JUN-2000; 2000JP-0191088.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Kurokawa T, Yamada T, Morimoto S;

DR WPI: 2002-139791/18.  
 XX N-PSDB: ABA99159, ABA99160.

PT Phosphatonin of human origin and DNA encoding it for diagnosis and  
 PT treatment of diseases associated with disorders of phosphate  
 PT metabolism, e.g., hyperphosphatemia, arteriosclerosis, heart failure,  
 PT diabetic renal disease and kidney failure

PS Claim 1: Fig 1-4; 130pp; Japanese.

XX This invention relates to human phosphatonin which has  
 CC the activity of lowering blood phosphate and increasing urinary  
 CC phosphate. The proteins are cytostatic, antidiabetic and  
 CC antiinflammatory in their action. Phosphatonin down-regulates  
 CC sodium-dependent phosphate transport in the kidney, it down-regulates  
 CC 25-hydroxy-vitamin D3-1alpha-hydroxylase in the kidney and up-regulates  
 CC 25-hydroxy-vitamin D3-24-hydroxylase in the kidney. Phosphatonin is  
 CC useful in the diagnosis, treatment and prevention of phosphate  
 CC metabolism related diseases such as hyperphosphatemia, arteriosclerosis,  
 CC heart failure, diabetic renal disease, kidney failure, acute coronary  
 CC disease and cystic fibrosis. This sequence represents human  
 CC phosphatonin.

XX Sequence 525 AA;

Query Match 100.0%; Score 126; DB 23; Length 525;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-11;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDLQERGDNDISPFSGDGPFFKD 23  
 DB 242 TDLQERGDNDISPFSGDGPFFKD 264

RESULT 11

AAE13227 standard; Protein: 525 AA.

AC AAE13227;

DT 12-FEB-2002 (first entry)

DE		Human oncogenic osteomalacia-related protein-1 (OOM-1).
KW	Human; oncogenic osteomalacia-related protein 1; OOM1; gene therapy;	
KM	phosphate homeostasis-related disease; rickets; osteomalacia; cardiac;	
KW	rhabdomyolysis; cardiomyopathy; tumoural calcinosis; renal failure;	
KX	bone mineralisation; phosphaturia; cellular process.	
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Peptide	1..16
FT		/label= Signal_peptide
FT	Protein	17..525
FT		(note= "Mature oncogenic osteomalacia-related protein-1"
FT	Domain	130..142
FT		/note= "Calcium binding motif"
FT	Domain	235..258
FT		/note= "Calcium binding motif"
FT	Region	247..249
FT		/note= "Cell attachment site"
FT	Domain	264..275
FT		/note= "Calcium binding motif"
FT	Domain	412..424
FT		/note= "Calcium binding motif"
FT	Modified-site	442..455
FT		/note= "Potential PKA phosphorylation site"
FT	Modified-site	477..481
FT		/note= "Glycosylation site"
FT	Domain	500..503
FT		/note= "Bacterial extracellular solute-binding protein motif"
PN	WO200172826-A2.	
XX		
PD	04-OCT-2001.	
XX		
PF	22-MAR-2001; 2001WO-US09289.	
XX		
PR	24-MAR-2000; 2000US-191786P.	
PR	19-OCT-2000; 2000US-241598P.	
XX		
PA	(GENZ ) GENZYME CORP.	
XX	(UTJO ) UNIV JOHN HOPKINS.	
PI	Schlavl S, Madden S, Manavalan P, Levine MD, Jan De Beur S;	
DR	WPJ: 2002-010720/O1.	
XX		
PT	New polynucleotide encoding a polypeptide useful for identifying	
PT	polynucleotide expression or polypeptide actively modulators used for	
PT	treating oncogenic osteomalacia, comprises the oncogenic	
PT	osteomalacia-related gene -	
XX		
PS	Claim 2; Fig 2; 65pp; English.	
CC	The invention relates to oncogenic osteomalacia-related protein 1	
CC	designated OOM1 (oncogenic osteomalacia-related factor) and its DNA	
CC	molecule. OOM1 protein is useful for treating a phosphate homeostasis-	
CC	related disease such as X-linked hypophosphataemia rickets, oncogenic	
CC	osteomalacia, rhabdomyolysis, cardiomyopathy, tumoural calcinosis and	
CC	renal failure. OOM1 proteins and DNA's are useful for modulating the	
CC	phenotype of a neoplastic cell associated with oncogenic osteomalacia;	
CC	modulating bone mineralisation; treating a disease characterised by	
CC	abnormal bone mineralisation; and treating an oncogenic osteomalacia-	
CC	associated symptom such as hypophosphatemia, phosphaturia, low serum	
CC	concentrations of 1,25-dihydroxy vitamin D or osteomalacia. OOM1 is	
CC	useful for modulating renal phosphate transport, which involves altering	
CC	the activity of OOM1 within a cell and then administering serine protease	
CC	11 to the cell. OOM1 DNA is useful for generating non-human transgenic	
CC	animal models, and for searching and identifying single nucleotide	
CC	polymorphisms which are mutants, variants of the gene in human	
CC	population. OOM1 is useful as an immunogen to produce antibodies against	

CC OOM1 and in vitro assays to screen for agents or drugs, which modulate  
CC cellular processes. The present sequence is human oncogenic  
CC osteomalacia-related protein-1 (OOM-1).  
CC Note: The present sequence states that this sequence is encoded by the  
CC DNA sequence shown in SEQ ID NO: 1 (AAD21889) of the specification.  
CC However this does not appear to be the case.  
XX

XX Sequence 525 AA:

XX

XX Query Match 100.0%; Score 126; DB 23; Length 525;  
XX Best Local Similarity 100.0%; Pred. No. 6,3e-11;  
XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

XX

XX 1 TDQERGDNDISPFSGDGPFXD 23  
XX |||||||||||||||||||  
XX 242 TDQERGDNDISPFSGDGPFXD 264

XX

XX RESULT 12  
XX AAB82923  
XX ID AAB82923 standard; Protein: 540 AA.  
XX AC AAB82923;  
XX XX  
XX DT 21-DEC-2001 (first entry)  
XX XX  
XX DE Human osteoregulin (mature polypeptide).  
XX XX  
XX KM Osteoregulin; human; bone; homeostasis; adipose; calcification;  
XX KM atherosclerosis; osteoporosis; osteopathic; antiarteriosclerotic;  
XX KM therapy.  
XX XX  
XX OS Homo sapiens.  
XX XX  
XX PN EPI130098-A2.  
XX PD  
XX PD 05-SEP-2001.  
XX XX  
XX PF 27-FEB-2001; 2001EP-0301768.  
XX XX  
XX PR 29-FEB-2000; 2000US-185617P.  
XX PR 22-SEP-2000; 2000US-234500P.  
XX XX  
XX PA (PE12 ) PFIZER PROD INC.  
XX XX  
XX PI Brown TA, De Wet JR, Gowen LC, Hames LM;  
XX XX  
XX DR WPI; 2001-604111/69.  
XX DR N-PSDB: AAB26811.

XX

XX Novel osteoregulin polypeptide useful for regulating bone homeostasis,  
XX PT adiposity and calcification of atherosclerotic plaques comprises  
XX PT measuring the activity of osteoregulin .  
XX

XX Claim 1; Page 59-61; 90pp; English.

XX

XX The present sequence is that of human osteoregulin mature  
XX polypeptide, i.e. lacking an N-terminal signal sequence.  
XX Osteoregulin is a novel protein which plays a role in regulating  
XX bone homeostasis, adiposity, and the calcification of  
XX atherosclerotic plaques. 2 Splice variants of human osteoregulin  
XX were identified (see also AAB82922). The invention provides novel  
XX osteoregulin proteins, nucleic acids which encode them, vectors,  
XX and antibodies, host cells which express heterologous osteoregulins, and  
XX animal cells and mammals with a targeted disruption of an  
XX osteoregulin gene. The invention also provides screening assays  
XX to identify modulators of osteoregulin activity as well as methods  
XX of treating mammals for diseases or disorders associated with  
XX osteoregulin activity. The modulators of activity may be useful  
XX in the manufacture of a medicament for, as well as for treating, a  
XX mammal in need of regulation of bone mass and/or density, adiposity,  
XX vascular flexibility, and/or atherosclerotic plaque calcification  
XX (claimed), for treating and preventing osteoporosis, and for

CC stimulating bone repair and regeneration.  
 XX Sequence 540 AA:  
 Query Match 100.0%; Score 126; DB 22; Length 540;  
 Best Local Similarity 100.0%; Pred. No. 6.5e-11;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDLOERGDNDISPFSGDQPFKD 23  
 |||||  
 Db 257 TDLOERGDNDISPFSGDQPFKD 279

## RESULT 13

AAB82921  
 ID AAB82921 standard; Protein; 556 AA.

XX AAB82921;

DT 21-DEC-2001 (first entry)

XX Human osteoregulin.

DE Osteoregulin; human; bone; homeostasis; adipose; calcification;

KW atherosclerosis; osteoporosis; osteopathic; arteriosclerotic;

XX Homo sapiens.

OS Key Location/Qualifiers

FT Peptide 1..37 /label= Signal\_peptide

FT Protein 38..556 /label= Mature\_protein

FT Misc-difference 57 /note= "encoded by TCA"

FT EPI130098-A2.

PD 05-SEP-2001.

PF 27-FEB-2001; 2001EP-0301768.

PR 29-FEB-2000; 2000US-185617P.

PR 22-SEP-2000; 2000US-234500P.

PA (PF12 ) PFIZER PROD INC.

PI Brown TA, De Wet JR, Gowen LC, Hames LM;

DR MPI: 2001-604111/69.

DR N-PSDB; AAH26809.

PT Novel osteoregulin polypeptide useful for regulating bone homeostasis,

PS measuring the activity of osteoregulin.

PS Claim 1; Page 48-49; 90pp; English.

CC The present sequence is that of human osteoregulin, a novel protein  
 CC which plays a role in regulating bone homeostasis, adiposity, and  
 CC the calcification of atherosclerotic plaques. The sequence is  
 CC predicted from the nucleotide sequence (see AAH26809) of isolated  
 CC osteoblast cDNA. A splice variant of human osteoregulin was also  
 CC identified (see AAB82920). The invention provides novel  
 CC osteoregulin proteins, nucleic acids which encode them, vectors,  
 CC antibodies, host cells which express heterologous osteoregulins, and  
 CC animal cells and mammals with a targeted disruption of an  
 CC osteoregulin gene. The invention also provides screening assays  
 CC to identify modulators of osteoregulin activity as well as methods  
 CC of treating mammals for diseases or disorders associated with  
 CC osteoregulin activity. The modulators of activity may be useful  
 CC in the manufacture of a medicament for, as well as for treating, a

CC mammal in need of regulation of bone mass and/or density, adiposity,  
 CC vascular flexibility, and/or atherosclerotic plaque calcification  
 CC (claimed), for treating and preventing osteoporosis, and for  
 CC stimulating bone repair and regeneration.

XX Sequence 556 AA;

Query Match 100.0%; Score 126; DB 22; Length 556;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-11;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDLOERGDNDISPFSGDQPFKD 23  
 |||||  
 Db 273 TDLOERGDNDISPFSGDQPFKD 295

## RESULT 14

AAU93703  
 ID AAU93703 standard; Peptide; 40 AA.

XX AAU93703;

DT 02-JUL-2002 (first entry)

DE Dental product bone growth enhancing peptide #23.

KW Dental product; toothpaste; mouthwash; dental floss; bone growth;

KW Integrin binding motif; RGD; skeletal disease; dental disease; tooth;

KW alveolar bone growth; osteoblast; odontoblast; osteopathic.

XX Synthetic.

OS PN NC0200213775-A1.

PD 21-FEB-2002.

PF 09-AUG-2001; 2001MO-US25101.

PR 16-AUG-2000; 2000US-225879P.

PA (BIGB-) BIG BEAR BIO INC.

PI Yoneda T, Nomizu M, Kumagai Y;

DR MPI: 2002-329525/36.

PT Dental product useful for treating skeletal diseases e.g. dental

PT diseases comprises a base material and a compound comprising specific

PT amino acid in a sequence containing the integrin binding motif

PS Claim 6; Page 20; 44pp; English.

XX The present invention relates to dental products such as toothpastes,

CC mouthwash and dental floss comprising a base material and a compound

CC which promotes bone growth. Such compounds are peptide sequences

CC comprising 10-50 amino acids and containing an integrin binding

CC motif such as RGD in the D- or L- form, preferably the L-configuration.

CC The peptides of the invention are useful for treating or preventing

CC skeletal diseases such as dental disease. The peptides enhance tooth

CC and/or alveolar bone growth on areas where deterioration has occurred,

CC as well as the growth or recruiting of osteoblast or odontoblast cells  
 CC on the surface of the new skeletal growth. AAU93681-AAU93726 represent  
 CC bone growth enhancing peptides for use in dental products.

XX Sequence 40 AA;

Query Match 80.2%; Score 101; DB 23; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-08;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 RGDNDISPFSGDQPFKD 23  
 |||||

Db 1 RGDNDISPFSGDQPFKD 18



## RESULT 15

AAO20353 standard: protein; 40 AA.

AAO20353:

31-MAY-2002 (first entry)

Protein of matrix extracellular phosphoglycoprotein containing RGD #23.

Bone growth; RGD motif; integrin binding motif; calcium binding motif;  
 glycosaminoglycan binding motif; bone loss; renal phosphate excretion;  
 alveolar; teeth; odontoblast; osteoclast; dental tissue; skeletal loss;  
 weakness; matrix extracellular phosphoglycoprotein.

Unidentified.

MO200214360-A1.

21-FEB-2002.

14-AUG-2001: 2001WO-US25542.

16-AUG-2000: 2000US-0641034.

19-MAR-2001: 2001US-0812485.

(BIGB-) BIG BEAR BIO INC.

Kumagai Y, Blacher RW, Yoneda T;

WPI: 2002-291971/33.

New peptide compound useful for reducing bone loss, is capable of  
 enhancing bone growth, and comprises an integrin binding motif,  
 glycosaminoglycan binding motif or a calcium binding motif -

Disclosure: Page 12: 50pp; English.

The invention relates to a peptide compound capable of enhancing bone  
 growth, and comprising 10-50 amino acids in a sequence, where the amino  
 acids are in D- or L- conformation and the sequence comprises a motif  
 selected from an integrin binding motif, a glycosaminoglycan binding  
 motif and a calcium binding motif. The peptide of the invention is useful  
 for reducing bone loss and for reducing renal phosphate excretion in an  
 individual. The peptide is useful for promoting regeneration of alveolar  
 bone and/or teeth, and increases the number and activity of osteoclasts  
 /osteoclasts that help form dental tissues. The peptide is also useful  
 for treating or preventing a condition associated with skeletal loss or  
 weakness. This sequence represents a protein of a matrix extracellular  
 phosphoglycoprotein containing an RGD motif of the invention.

Sequence 40 AA:

Query Match 80.2%; Score 101; DB 23; Length 40;

Best Local Similarity 100.0%; Pred. No. 2.4e-08;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RGDNDISPFSGDGPFPKD 23  
 |||||  
 DB 1 RGDNDISPFSGDGPFPKD 18

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 13:16:42 : Search time 14 Seconds  
(without alignments)  
48.338 Million cell updates/sec

Title: US-09-641-034-47

Perfect score: 126

Sequence: 1 TDQERGDNDISFGDQPFKD 23

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 segs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents, AA: \*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep: \*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep: \*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep: \*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep: \*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS.COMB.pep: \*  
6: /cgn2\_6/ptodata/1/1aa/Backfilest1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	41.3	301	2	US-08-355-844-2
2	52	41.3	301	5	PCT-US95-16126-2
3	47	37.3	815	4	US-09-134-001C-3614
4	47	37.3	1198	4	US-09-199-637A-405
5	45.5	36.1	96	2	US-08-341-843B-9
6	45.5	36.1	96	2	US-08-427-497E-14
7	45.5	36.1	1260	4	US-08-506-296B-21
8	44	34.9	348	1	US-08-176-126B-2
9	44	34.9	348	1	US-08-669-435-2
10	44	34.9	348	5	PCT-US94-14431A-2
11	44	34.9	475	1	US-08-278-091-6
12	44	34.9	475	1	US-08-483-859-6
13	44	34.9	475	1	US-08-472-173-6
14	44	34.9	475	1	US-08-350-741-2
15	44	34.9	475	1	US-08-487-167-6
16	44	34.9	475	2	US-08-482-816-6
17	44	34.9	475	2	US-08-296-149-6
18	44	34.9	475	2	US-08-801-499-6
19	44	34.9	475	2	US-08-463-875A-2
20	44	34.9	475	2	US-08-615-271-6
21	44	34.9	475	3	US-09-074-660-6
22	44	34.9	475	3	US-09-074-659-6
23	44	34.9	475	3	US-09-106-468-6
24	44	34.9	475	4	US-09-106-466A-6
25	44	34.9	475	4	US-09-106-467-6
26	43	34.1	169	3	US-08-476-509B-28
27	42.5	33.7	306	1	US-08-330-978-1

28	42.5	33.7	306	1	US-08-474-042-1	Sequence 1, Appl1
29	42.5	33.7	306	1	US-08-484-558-1	Sequence 1, Appl1
30	42.5	33.7	306	1	US-08-774-592-1	Sequence 1, Appl1
31	42.5	33.7	448	1	US-08-295-411-3	Sequence 3, Appl1
32	42.5	33.7	448	2	US-08-955-471-3	Sequence 3, Appl1
33	42.5	33.7	448	5	PCT-US92-10068-1	Sequence 1, Appl1
34	42.5	33.7	448	5	PCT-US92-10242-3	Sequence 3, Appl1
35	42.5	33.7	488	1	US-08-487-037-1	Sequence 1, Appl1
36	42.5	33.7	545	2	US-08-990-114-1	Sequence 1, Appl1
37	42.5	33.7	545	4	US-09-241-333-1	Sequence 1, Appl1
38	41.5	32.9	364	4	US-08-858-207A-313	Sequence 13, App
39	41.5	32.9	811	4	US-08-983-275-2	Sequence 2, Appl1
40	40.5	32.1	397	4	US-09-087-134-17	Sequence 17, Appl1
41	40	31.7	469	2	US-08-416-870C-6	Sequence 6, Appl1
42	40	31.7	795	4	US-09-370-807-6	Sequence 6, Appl1
43	40	31.7	795	4	US-09-921-259-6	Sequence 6, Appl1
44	40	31.7	831	2	US-09-047-026A-4	Sequence 4, Appl1
45	40	31.7	912	5	PCT-US95-03747-2	Sequence 2, Appl1

## ALIGNMENTS

RESULT 1  
US-08-355-844-2  
Sequence 2, Application US/08355844  
Patent No. 5940307  
GENERAL INFORMATION:  
APPLICANT: Fischbarg, Jorge  
APPLICANT: Czelegiedy, Ferenc  
APPLICANT: Isevorich, Pavel  
APPLICANT: Li, Jun  
TITLE OF INVENTION: A METHOD FOR PREDICTING PROTEIN  
NUMBER OF INVENTIONS: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10112-0228  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/355,844  
FILING DATE: 14-DEC-1994  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Tang, Henry Y.S.  
REGISTRATION NUMBER: 29,705  
REFERENCE/DOCKET NUMBER: A29927-50/29910  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-408-2586  
TELEFAX: 212-765-2519  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 301 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORGANISM: Rhodobacter capsulatus  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..301  
OTHER INFORMATION: Porin protein  
US-08-355-844-2

Query Match 41.3%; Score 52; DB 2; Length 301;  
Best Local Similarity 58.8%; Pred. No. 2.8;  
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
OY 1 TDLQERGDNDISFSGD 17  
111:1111111111  
Db 92 TDLDRGNDIPYLTGD 108

## RESULT 2

PCT-US95-16126-2  
Sequence 2 Application PC/TUS9516126  
GENERAL INFORMATION:  
APPLICANT: Fischbarg, Jorge  
APPLICANT: Czelegedy, Ferenc  
APPLICANT: Iserovich, Pavel  
APPLICANT: Li, Jun  
APPLICANT: Cheung, Min  
TITLE OF INVENTION: A METHOD FOR PREDICTING PROTEIN  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10112-0228  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/16126  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/355,844  
FILING DATE: 14-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Tang, Henry T.S.  
REGISTRATION NUMBER: 29,705  
REFERENCE/DOCKET NUMBER: A29927-50/29910  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-408-2586  
TELEFAX: 212-765-2519  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 301 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Rhodobacter capsulatus  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..301  
OTHER INFORMATION: Porin protein  
PCT-US95-16126-2

Query Match 41.3%; Score 52; DB 5; Length 301;  
Best Local Similarity 58.8%; Pred. No. 2.8;  
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
OY 1 TDLQERGDNDISFSGD 17  
111:1111111111  
Db 92 TDLDRGNDIPYLTGD 108

## RESULT 3

US-09-134-001C-3614  
Sequence 3614, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
PRIOR FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 3614  
LENGTH: 815  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3614

Query Match 37.3%; Score 47; DB 4; Length 815;  
Best Local Similarity 58.8%; Pred. No. 50;  
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 7 GDNDISFSGDGPFRD 23  
111:1111111111

Db 481 GDEDLPSVGPQGVFRD 497

## RESULT 4

US-09-199-637A-405  
Sequence 405, Application US/09199637A  
Patent No. 6355411  
GENERAL INFORMATION:  
APPLICANT: Ausubel, Frederick  
APPLICANT: Goodman, Howard M.  
APPLICANT: Rahme, Laurence G.  
APPLICANT: Mahajan-Miklos, Shailina  
APPLICANT: Tan, Man-Wah  
APPLICANT: Drenkard, Eliana  
APPLICANT: Tsongalis, John  
TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID  
FILE REFERENCE: 00786/361002  
CURRENT APPLICATION NUMBER: US/09/199,637A  
PRIOR FILING DATE: 1998-11-25  
PRIOR APPLICATION NUMBER: 60/066,517  
PRIOR FILING DATE: 1997-11-25  
NUMBER OF SEQ ID NOS: 437  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 405  
LENGTH: 1198  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-199-637A-405

Query Match 37.3%; Score 47; DB 4; Length 1198;  
Best Local Similarity 52.4%; Pred. No. 77;  
Matches 11; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

OY 1 TDLQERGDNDISFSGD 17  
111:1111111111  
Db 400 TDLRPSIDNLRPFMAYKSGD 420

## RESULT 5

US-08-341-843B-9  
Sequence 9, Application US/08341843B  
Patent No. 5872225  
GENERAL INFORMATION:  
APPLICANT: Lemmon, Vance

TITLE OF INVENTION: A Method for Characterizing the  
 Nucleotide Sequence of L1CAM and  
 Patent No. 5872225  
 TITLE OF INVENTION: the Nucleotide Sequence  
 Characterized Thereby  
 NUMBER OF SEQUENCES: 39  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Fay, Sharpe, Beall, Fagan,  
 ADDRESS: Minnich & McKee  
 STREET: 1100 Superior Avenue  
 STREET: Suite 700  
 CITY: Cleveland  
 STATE: Ohio  
 COUNTRY: U.S.A.  
 ZIP: 44114-2518  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 720 KB  
 MEDIUM TYPE: Storable  
 COMPUTER: Compaq Prolinea 5100e  
 OPERATING SYSTEM: DOS 5.0  
 SOFTWARE: ASCII  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/341,843B  
 FILING DATE: No. 5872225ember 18, 1994  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 07/904,991  
 FILING DATE: June 26, 1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Minnich, Richard J.  
 REGISTRATION NUMBER: 24,175  
 REFERENCE/DOCKET NUMBER: CWR 2 149-1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (216) 861-5582  
 TELEFAX: (216) 241-1666  
 TELEX: (216) 980162  
 INFORMATION FOR SEQ ID NO: 9:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 96  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: amino acids  
 HYPOTHETICAL: irrelevant.  
 ANTI-SENSE: no  
 ORIGINAL SOURCE:  
 ORGANISM: mouse  
 INDIVIDUAL ISOLATE: 8 day old mouse brain  
 IMMEDIATE SOURCE:  
 LIBRARY: lambda GT 10 and lambda GT11  
 CLONE: synthesis of several clones  
 PUBLICATION INFORMATION:  
 AUTHORS: Moos, M.  
 AUTHORS: Tacke, R.  
 AUTHORS: Scherer, H.  
 AUTHORS: Teplov, D.  
 AUTHORS: Frub, K.  
 AUTHORS: Schachner, M.  
 TITLE: Neural adhesion molecule L1 is a  
 member of the immunoglobulin  
 TITLE: superfamily with binding domains  
 TITLE: similar to fibronectin  
 JOURNAL: NATURE  
 VOLUME: 334  
 ISSUE:  
 PAGES: 701-703  
 DATE: 1988  
 15-08-341-843B-9

```

OY      2 DLOERGNDISPFGDQ 19
         |||||:|  |  |
Db      41 DLOERGDSD-KYFIEDGK 57

```

RESULT 6  
US-08-427-497E-14

sequence 14, Application US/08427497E  
Patent No. 5969124

GENERAL INFORMATION

APPLICANT: Lemmon, Vance

TITLE OF INVENTION:	A Method for Characterizing the
TITLE OF INVENTION:	Nucleotide Sequence of LICAM and
Patent No. 5969124	

```

; TITLE OF INVENTION: the Nucleotide Sequence
; TITLE OF INVENTION: Characterized Thereby
; NUMBER OF SEQUENCES: 44
;

```

ADDRESSEE: Fay, Sharpe, Beall, Fagan  
ADDRESSEE: Mlinich & McKee

```

; STREET: 1100 Superior Avenue
; STREET: Suite 700
; CITY: Cleveland
;

```

STATE: Ohio  
COUNTRY: U.S.A.  
ZIP: 44114-2518

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3 50 inch 720 kb

```

```

' MEDIUM AFFE: DISKETTES, 5.25 INCH
: MEDIUM TYPE: STORABLE

```

COMPUTER: Compaq ProLinea

```
OPERATING SYSTEM
SOFTWARE: ASCII
```

; CURRENT APPLICATION DATA:

APPLICATION NUMBER:  
; FILING DATE: APR 11

CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/004 001

DECLASSIFICATION NUMBER: 01/3  
FILING DATE: June 26, 19

ATTORNEY/AGENT INFORMATION:

NAME: MINNICH, RICHARD J  
REGISTRATION NUMBER: 24.

REFERENCE/DOCKET NUMBER: CWR 2 149-3-1

TELECOMMUNICATION INFORMATION  
TELEPHONE: (216) 861-5583

TELEFAX: (216) 241-1666

TELEX: (216) 980162

SEQUENCE CHARACTERISTICS: 14

LENGTH: 9

TYPE: amino acid  
STRANDEDNESS: single

TOPLOGY: linear

MOLECULE TYPE: amino acids

ANTI-SENSE: no

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE. 8 day old mouse brain

IMMEDIATE SOURCE ;

LIBRARY: lamda GT 10 and lamda GT11  
CLONE: synthetis of several clones

PUBLICATION INFORMATION:

AUTHORS: MOOS, I

AUTHORS: Jacke, R.  
AUTHORS: Scherer, H.

**AUTHORS:** Teplov, D.

AUTHORS: Fruh, K.  
: Schachner M

TITLE: Neural adhesion molecule L1 is a

TITLE: member of the immunoglobulin

```

; TITLE: similar to fibronectin
; IILE: superfamily with binding domains

```

; JOURNAL: NATURE  
 ; VOLUME: 334  
 ; ISSUE:  
 ; PAGES: 701-703  
 ; DATE: 1988  
 ;  
 US-08-427-497E-14

Query Match	36.18;	Score 45.5;	DB 2;	Length 96;
Best Local Similarity	61.10;	Pred. NO. 7.3;		
Matches 11; Conservative	2;	Mismatches 4;	Indels 1;	Gaps 1

```

QY      2 DLQERGNDISPFGDGQ 19
          |||||:|  | |
DB      41 DLQERGDS-KYFIEDGK 57

```

RESULT 7  
US-08-506-296B-21  
: Sequence 21, Application US/08506296B  
: Patent No. 6313265

1 APPLICANT: Phillips, Greg  
 2 APPLICANT: Cunningham, Bruce A.  
 3 APPLICANT: Crossin, Kathryn L.  
 4 TITLE OF INVENTION: NEURITE OUTGROWTH-PROMOTING POLYPEPTIDES  
 5 TITLE OF INVENTION: CONTAINING FIBRONECTIN TYPE III REPEATS AND METHODS OF USE  
 6 NUMBER OF SEQUENCES: 77  
 7 CORRESPONDENCE ADDRESS:

ADDRESSEE: The Scripps Research Institute  
STREET: 10550 No. 6313265th Torrey Pines Road, TPC-8  
CITY: La Jolla  
STATE: California  
COUNTRY: U.S.

```

1  COMPUTER READABLE FORM:
2  MEDIUM TYPE: Floppy disk
3  COMPUTER: IBM PC compatible
4  OPERATING SYSTEM: PC-DOS/MS-DOS
5  SOFTWARE: Patentin Release #1.0, Version #1.25
6  CURRENT APPLICATION DATA:
7  APPLICATION NUMBER: US/08/506,296B
8  FILING DATE: 24-JUL-1995
9  CLASSIFICATION: 514

```

NAME: Pittling, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI 488.C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 554-2937

```

SEQUENCE CHARACTERISTICS:
LENGTH: 1260 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-506-296B-21

```

Query Match	36.18;	Score 45.5;	DB 4;	Length 1260;
Best Local Similarity	61.18;	Pred. No. 1.4e+02;		
Matches 11; Conservative	2;	Mismatches 4;	Indels 1;	Gaps 1;

```

QY      2 DLQERGNDISPFSGDQ 19
          |||||:|  ||:
Db      558 DLQERGSD-KYFIEDGK 574

```

RESULT 8  
US-08-176-126B-2  
; Sequence 2, Application US/08176126B  
; Patent No. 5589358  
; GENERAL INFORMATION:  
; APPLICANT: Dawson, Paul A.

TITLE OF INVENTION: 116AL BILE ACID TRANSPORTER COMPOSITIONS AND  
 METHOD OF INVENTION: METHODS  
 NUMBER OF SEQUENCES: 5  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Arnold, White & Durkee  
 STREET: P. O. Box 4433  
 CITY: Houston, Texas 77240-4433

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: 100-475-1000  
IDENTIFICATION NUMBER: 100-475-1000

APPLICATION NUMBER: US/08/176,126B  
FILING DATE: 29-DEC-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Parker, David L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: WAKE:002/PAR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577

```

: INFORMATION FOR SEQ ID NO: 2:
:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 348 amino acids
:   TYPE: amino acid
:   TOPOLOGY: linear
:   MOLECULE TYPE: protein
:
US-08-176-126B-2

```

Query Match	34.9%	Score 44	DB 1	Length 348
Best Local Similarity	53.3%	Pred. No. 53		
Matches	8	Conservative	4	Mismatches 3
				Indels 0
				Gaps 0

```
QY      1 TDLQERGDNDISPFS 15
          | : | | : | : | |
Db      320 TELQEKTDNEMEPFS 334
```

RESULT 9  
US-08-669-435-2  
; Sequence 2, Application US/08669435  
; Patent No. 5869265  
; GENERAL INFORMATION:

APPLICANT: Dawson, Paul A.  
TITLE OF INVENTION: LIPAL BILE ACID TRANSPORTER COMPOSITIONS AND  
METHODS  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P. O. Box 4433

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/669,435  
FILING DATE: 26-JUN-1996  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/176,126  
FILING DATE: 29-DEC-1993  
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:  
NAME: PARKER, DAVID L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: MAKE:002/PAR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577  
TELEX: na  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 348 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-669-435-2

Query Match 34.9% Score 44; DB 2; Length 348;  
Best Local Similarity 53.3%; Pred. No. 53;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 TDQERGDNDISFES 15  
1:111:11:11  
Db 320 TELQEKTDNEMEPERS 334

RESULT 10  
PCT-US94-14431A-2  
Sequence 2, Application PC/TUS9414431A  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: IDEAL BILE ACID TRANSPORTER COMPOSITIONS  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P. O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: United States of America  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
SOFTWARE: Patent Release #1.0, Version  
SOFTWARE: #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/14431A  
FILING DATE: 29-DEC-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/176,126  
FILING DATE: 29-DEC-1993  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: PARKER, DAVID L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: MAKE005P--  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
TELEX: 79-0924(1) GENERAL INFORMATION:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 348 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US94-14431A-2

Query Match 34.9% Score 44; DB 5; Length 348;  
Best Local Similarity 53.3%; Pred. No. 53;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 TDQERGDNDISFES 15  
1:111:11:11  
Db 320 TELQEKTDNEMEPERS 334

RESULT 11  
US-08-278-091-6  
Sequence 6, Application US/08278091  
Patent No. 5506139  
GENERAL INFORMATION:  
APPLICANT: LOOSMORE, Sheena M  
APPLICANT: YANG, Yan-Ping  
APPLICANT: CHONG, Pele  
APPLICANT: OOMEN, Raymond P.  
APPLICANT: KLEIN, Michel H.  
TITLE OF INVENTION: Analog of Haemophilus Hln47 Protein with  
REDUCED PROTEASE ACTIVITY  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Slim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/278,091  
FILING DATE: 21-JUL-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-371  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 475 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-278-091-6

Query Match 34.9% Score 44; DB 1; Length 475;  
Best Local Similarity 57.1%; Pred. No. 76;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 10 DISFGDGOPEKD 23  
1:111:11:11  
Db 79 DQSPFCODSSPQN 92

RESULT 12  
US-08-483-859-6  
Sequence 6, Application US/08483859  
Patent No. 5656436  
GENERAL INFORMATION:  
APPLICANT: LOOSMORE, Sheena M.  
APPLICANT: YANG, Yan-Ping  
APPLICANT: CHONG, Pele  
APPLICANT: OOMEN, Raymond P.  
APPLICANT: KLEIN, Michel H.  
TITLE OF INVENTION: Analog of Haemophilus Hln47 Protein with  
REDUCED PROTEASE ACTIVITY  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Slim & McBurney

STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/483,859  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/296,149  
FILING DATE: 26-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/278,091  
FILING DATE: 21-JUL-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-495 MIS.vg  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 475 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-483-859-6

Query Match 34.9%; Score 44; DB 1; Length 475;  
Best Local Similarity 57.1%; Pred. No. 76;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 10 DISFSGDGPFD 23  
DB 79 DDSFPCDGSFON 92

RESULT 13  
US-08-472-173-6  
Sequence 6, Application US/08472173  
Patent No. 565353  
GENERAL INFORMATION:  
APPLICANT: LOOSMORE, Sheena M  
APPLICANT: YANG, Yan-Ping  
APPLICANT: CHONG, Pele  
APPLICANT: COHEN, Raymond P.  
APPLICANT: KLEIN, Michel H.  
TITLE OF INVENTION: Analog of Haemophilus Hln47 Protein with  
TITLE OF INVENTION: Reduced Protease Activity  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,173  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/296,149  
FILING DATE: 26-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/278,091  
FILING DATE: 21-JUL-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-493 MIS.vg  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 475 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-472-173-6

Query Match 34.9%; Score 44; DB 1; Length 475;  
Best Local Similarity 57.1%; Pred. No. 76;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 10 DISFSGDGPFD 23  
DB 79 DDSFPCDGSFON 92

RESULT 14  
US-08-350-741-2  
Sequence 2, Application US/08350741  
Patent No. 5804194  
GENERAL INFORMATION:  
APPLICANT: DOUGAN G.  
APPLICANT: CHARLES I.G.  
APPLICANT: HORMACHE C.E.  
APPLICANT: JOHNSON K.S.  
APPLICANT: CHATFIELD S.N.  
TITLE OF INVENTION: LIVE VACCINES  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON and VANDERHAYE PC  
STREET: 8th FLOOR, 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: USA  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/350,741  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/952,737  
FILING DATE: 09-MAY-1994  
APPLICATION NUMBER: US 07/952,737  
FILING DATE: 30-NOV-1992  
APPLICATION NUMBER: GB 9007194.5  
FILING DATE: 30-MAR-1990  
APPLICATION NUMBER: PCT/GB91/00484  
FILING DATE: 28-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: WILSON, MARY J.  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 117-158  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000

```

: TELEFAX: (703) 816-4100
:
: TELEX: 200797 NIXN UR
:
: INFORMATION FOR SEQ ID NO:
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 475 amino acids
:
: TYPE: amino acid
:
: TOPOLOGY: linear
:
: MOLECULE TYPE: protein
:
US-08-350-741-2

```

Query Match	34.98;	Score 44;	DB 1;	Length 475;
Best Local Similarity	57.18;	Pred. No. 76;		
Matches	8; Conservative	2; Mismatches	4; Indels	

```
QY      10 DISPSGDGPFKD 23  
         1 111 11 11:1  
DB       79 DDSPFCQDCSPFQN 92
```

```

US-08-487-167-6
: Sequence 6, Application US/08487167
: Patent No. 5869302
:
: GENERAL INFORMATION:
: APPLICANT: LOOSMORE, Sheena M.
: APPLICANT: YANG, Yan-Ping
: APPLICANT: CHONG, Pele
: APPLICANT: COHEN, Raymond P.
: APPLICANT: KLEIN, Michel H.
: TITLE OF INVENTION: Analog of Haemophilus Hin47 protein with
: TITLE OF INVENTION: Reduced Protease Activity
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sim & McBurney
: STREET: Suite 701, 330 University Avenue
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5G 1R7
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/487,167
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/296,149
: FILING DATE: 26-AUG-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/278,091
: FILING DATE: 21-JUL-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Stewart, Michael I.
: REGISTRATION NUMBER: 24,973
: REFERENCE/DOCKET NUMBER: 1038-508 MIS:vg
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 595-1155
: TELEFAX: (416) 595-1163
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 475 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
US-08-487-167-6
Query Match 34.9% Score 44; DB 2; Length 475;
Best Local Similarity 57.1% Pred. NO. 76;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 13:18:07 ; Search time 10 Seconds  
(without alignments)  
45.721 Million cell updates/sec

Title: US-09-641-034-47

Perfect score: 126

Sequence: 1 TDLQERGDNDISFSGDGPFDK 23

Scoring table: BL0SUM62  
Gapop 10.0 , Gapext 0.5

Searched: 120991 seqs, 19878514 residues

120991

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications, AA.\*

1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	126	100.0	23	9	US-09-812-485A-49
2	126	100.0	97	9	US-09-812-485A-1
3	126	100.0	525	10	US-09-814-550-2
4	101	80.2	40	9	US-09-812-485A-23
5	86	68.3	19	10	US-09-814-550-6
6	84	66.7	15	9	US-09-812-485A-48
7	84	66.7	33	9	US-09-812-485A-48
8	82	65.1	15	9	US-09-812-485A-47
9	75	59.5	15	9	US-09-812-485A-46
10	73.5	58.3	40	9	US-09-812-485A-32
11	72.5	57.5	30	9	US-09-812-485A-26
12	72.5	57.5	35	9	US-09-812-485A-25
13	66	52.4	12	9	US-09-812-485A-42
14	57	45.2	40	9	US-09-812-485A-32
15	57	45.2	45	9	US-09-812-485A-27
16	53	42.1	1458	12	US-10-054-691-2
17	52	41.3	19	10	US-09-814-550-5
18	52	41.3	301	10	US-09-905-176-23
19	49	38.9	9	10	US-09-814-550-8

20	49	38.9	326	9	US-09-738-626-4959	Sequence 4959, App
21	47	37.3	751	10	US-09-815-242-5832	Sequence 5832, App
22	47	37.3	825	10	US-09-815-242-12963	Sequence 12963, App
23	46.5	36.9	1137	12	US-10-100-912-4	Sequence 437, App
24	46.5	36.9	1141	9	US-09-978-295A-437	Sequence 437, App
25	46.5	36.9	1141	9	US-09-978-697-437	Sequence 437, App
26	46.5	36.9	1141	9	US-09-978-192A-437	Sequence 437, App
27	46.5	36.9	1141	9	US-09-999-882A-437	Sequence 437, App
28	46.5	36.9	1141	9	US-09-978-189-437	Sequence 437, App
29	46.5	36.9	1141	9	US-10-174-590-120	Sequence 120, App
30	46.5	36.9	1141	12	US-10-052-586-120	Sequence 120, App
31	46.5	36.9	1141	10	US-09-815-242-11959	Sequence 11959, App
32	45.5	36.1	347	10	US-09-815-242-11959	Sequence 11959, App
33	45.5	36.1	1260	9	US-10-024-918-27	Sequence 27, App
34	44	34.9	9	10	US-09-814-550-7	Sequence 7, App
35	43.5	34.5	354	10	US-09-825-414-70	Sequence 70, App
36	43	34.1	1209	9	US-10-108-605-349	Sequence 349, App
37	43	34.1	2783	10	US-09-816-669A-14	Sequence 14, App
38	42.5	33.7	216	10	US-09-925-299-841	Sequence 841, App
39	42.5	33.7	545	10	US-09-878-242-1	Sequence 1, App
40	42.5	33.7	561	10	US-09-821-687-4	Sequence 4, App
41	42.5	33.7	579	10	US-09-925-300-1415	Sequence 1415, App
42	42	33.3	15	9	US-09-812-485A-45	Sequence 45, App
43	42	33.3	47	9	US-09-812-485A-2	Sequence 2, App
44	42	33.3	333	10	US-09-825-301-952	Sequence 952, App
45	41.5	32.9	639	10	US-09-815-242-5390	Sequence 5390, App

## ALIGNMENTS

RESULT 1  
US-09-812-485A-49  
Sequence 49, Application US/09812485A  
Publication No. US20020197267A1  
GENERAL INFORMATION:  
APPLICANT: Kumagai, Yoshinari  
APPLICANT: Blacher, Toshitaki  
TITLE OF INVENTION: Peptide Binding Motif Containing  
TITLE OF INVENTION: Peptides and Methods of Treating Skeletal Diseases  
FILE REFERENCE: BEAR-006CIP  
CURRENT APPLICATION NUMBER: US/09/812,485A  
CURRENT FILING DATE: 2001-03-19  
PRIOR APPLICATION NUMBER: 09/641,034  
PRIOR FILING DATE: 2000-08-16  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 49  
LENGTH: 23  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: D-00006 peptide  
NAME/KEY: AMIDATION  
LOCATION: 15  
US-09-812-485A-49

Query Match 100.0% Score 126; DB 9; Length 23;  
Best Local Similarity 100.0% Pred No. 5.2e-13;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDLQERGDNDISFSGDGPFDK 23  
DB 1 TDLQERGDNDISFSGDGPFDK 23

RESULT 2  
US-09-812-485A-1  
Sequence 1, Application US/09812485A  
Publication No. US20020197267A1  
GENERAL INFORMATION:  
APPLICANT: Kumagai, Yoshinari

GENERAL INFORMATION: yoshinari  
APPLICANT: Kumagai, yoshinari  
APPLICANT: Blacher, Russel  
APPLICANT: yoneda, toshiyuki  
TITLE OF INVENTION: Integrin Binding Molecule Containing  
TITLE OF INVENTION: Peptides and Methods of Treating Skeletal Diseases  
FILE REFERENCE: BEAR-006CIP  
CURRENT APPLICATION NUMBER: US/09/012,485A  
CURRENT FILING DATE: 2001-03-19  
PRIOR APPLICATION NUMBER: 09/641,034

```

; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: D-00005 peptide
; NAME/KEY: AMIDATION
; LOCATION: 15
US-09-812-485A-48

```

```

Query Match
Best Local Similarity 66.7%; Score 84; DB 9; Length 15;
Matches 15; Conservative 0; Pred. No. 6.7e-07; Mismatches 0; Indels 0; Gaps 0;
QY 9 NDSPFSGDGPDKD 23
DB 1 NDSPFSGDGPDKD 15

```

```

RESULT 7
US-09-812-485A-30
; Sequence 30, Application US/09812485A
; Publication No. US20020197267A1
; GENERAL INFORMATION:
; APPLICANT: Kumagai, Yoshinari
; APPLICANT: Blacher, Russel
; APPLICANT: Yoneda, Toshiyuki
; TITLE OF INVENTION: Integrin Binding Motif Containing
; FILE REFERENCE: BEAR-006CIP
; CURRENT APPLICATION NUMBER: US/09/812,485A
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 09/641,034
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptidic compound
US-09-812-485A-30

```

```

Query Match
Best Local Similarity 66.7%; Score 84; DB 9; Length 33;
Matches 15; Conservative 0; Pred. No. 1.6e-06; Mismatches 0; Indels 0; Gaps 0;
QY 9 NDSPFSGDGPDKD 23
DB 1 NDSPFSGDGPDKD 15

```

```

RESULT 8
US-09-812-485A-47
; Sequence 47, Application US/09812485A
; Publication No. US20020197267A1
; GENERAL INFORMATION:
; APPLICANT: Kumagai, Yoshinari
; APPLICANT: Blacher, Russel
; APPLICANT: Yoneda, Toshiyuki
; TITLE OF INVENTION: Integrin Binding Motif Containing
; FILE REFERENCE: BEAR-006CIP
; CURRENT APPLICATION NUMBER: US/09/812,485A
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 09/641,034
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 47
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: D-00004 peptide
; NAME/KEY: AMIDATION
; LOCATION: 15
US-09-812-485A-47

```

```

Query Match
Best Local Similarity 65.1%; Score 82; DB 9; Length 15;
Matches 15; Conservative 0; Pred. No. 1.3e-06; Mismatches 0; Indels 0; Gaps 0;
QY 5 ERGDNDISPSGDQ 19
DB 1 ERGDNDISPSGDQ 15

```

```

RESULT 9
US-09-812-485A-46
; Sequence 46, Application US/09812485A
; Publication No. US20020197267A1
; GENERAL INFORMATION:
; APPLICANT: Kumagai, Yoshinari
; APPLICANT: Blacher, Russel
; APPLICANT: Yoneda, Toshiyuki
; TITLE OF INVENTION: Integrin Binding Motif Containing
; FILE REFERENCE: BEAR-006CIP
; CURRENT APPLICATION NUMBER: US/09/812,485A
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 09/641,034
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: D-00003 peptide
; NAME/KEY: AMIDATION
; LOCATION: 15
US-09-812-485A-46

```

```

Query Match
Best Local Similarity 59.5%; Score 75; DB 9; Length 15;
Matches 14; Conservative 0; Pred. No. 1.5e-05; Mismatches 0; Indels 0; Gaps 0;
QY 1 TDLOERGDNDISPF 14
DB 2 TDLOERGDNDISPF 15

```

```

RESULT 10
US-09-812-485A-24
; Sequence 24, Application US/09812485A
; Publication No. US20020197267A1
; GENERAL INFORMATION:
; APPLICANT: Kumagai, Yoshinari
; APPLICANT: Blacher, Russel
; APPLICANT: Yoneda, Toshiyuki
; TITLE OF INVENTION: Integrin Binding Motif Containing
; FILE REFERENCE: BEAR-006CIP
; CURRENT APPLICATION NUMBER: US/09/812,485A
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 09/641,034
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24

```

LENGTH: 40  
TYPE: PRF  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: peptidic compound  
US-09-812-485A-24

Query Match 58.3%; Score 73.5; DB 9; Length 40;  
Best Local Similarity 83.3%; Pred. No. 7.7e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

OY 6 RGDNDISFGSGDGPFRD 23  
DB 4 RGD---SPFGSDGCPFRD 18

RESULT 11  
US-09-812-485A-26  
Sequence 26, Application US/09812485A  
Publication No. US20020197267A1  
GENERAL INFORMATION:  
APPLICANT: Kumagai, Yoshinari  
APPLICANT: Blacher, Russel  
APPLICANT: Yoneda, Toshiyuki  
TITLE OF INVENTION: Integrin Binding Motif Containing  
Peptides and Methods of Treating Skeletal Diseases  
FILE REFERENCE: BEAR-006CIP  
CURRENT APPLICATION NUMBER: US/09/812,485A  
CURRENT FILING DATE: 2001-03-19  
PRIOR APPLICATION NUMBER: 09/641,034  
PRIOR FILING DATE: 2000-08-16  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 26  
LENGTH: 30  
TYPE: PRF  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: peptidic compound  
US-09-812-485A-26

Query Match 57.5%; Score 72.5; DB 9; Length 30;  
Best Local Similarity 83.3%; Pred. No. 7.8e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

OY 9 NDISPFS---GDGCPFRD 23  
DB 1 NDISPFSGDRCGCPFRD 18

RESULT 12  
US-09-812-485A-25  
Sequence 25, Application US/09812485A  
Publication No. US20020197267A1  
GENERAL INFORMATION:  
APPLICANT: Kumagai, Yoshinari  
APPLICANT: Blacher, Russel  
APPLICANT: Yoneda, Toshiyuki  
TITLE OF INVENTION: Integrin Binding Motif Containing  
Peptides and Methods of Treating Skeletal Diseases  
FILE REFERENCE: BEAR-006CIP  
CURRENT APPLICATION NUMBER: US/09/812,485A  
CURRENT FILING DATE: 2001-03-19  
PRIOR APPLICATION NUMBER: 09/641,034  
PRIOR FILING DATE: 2000-08-16  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 25  
LENGTH: 35  
TYPE: PRF  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: peptidic compound

US-09-812-485A-25  
Query Match 57.5%; Score 72.5; DB 9; Length 35;  
Best Local Similarity 83.3%; Pred. No. 9.3e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

OY 9 NDISPFS---SGDGPFRD 23  
DB 1 NDISPFRDSDGCPFRD 18

RESULT 13  
US-09-812-485A-42  
Sequence 42, Application US/09812485A  
Publication No. US20020197267A1  
GENERAL INFORMATION:  
APPLICANT: Kumagai, Yoshinari  
APPLICANT: Blacher, Russel  
APPLICANT: Yoneda, Toshiyuki  
TITLE OF INVENTION: Integrin Binding Motif Containing  
Peptides and Methods of Treating Skeletal Diseases  
FILE REFERENCE: BEAR-006CIP  
CURRENT APPLICATION NUMBER: US/09/812,485A  
CURRENT FILING DATE: 2001-03-19  
PRIOR APPLICATION NUMBER: 09/641,034  
PRIOR FILING DATE: 2000-08-16  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 42  
LENGTH: 12  
TYPE: PRF  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: calcium binding motif  
US-09-812-485A-42

Query Match 52.4%; Score 66; DB 9; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00027;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 DNDISFGSDGQ 19  
DB 1 DNDISFGSDGQ 12

RESULT 14  
US-09-812-485A-32  
Sequence 32, Application US/09812485A  
Publication No. US20020197267A1  
GENERAL INFORMATION:  
APPLICANT: Kumagai, Yoshinari  
APPLICANT: Blacher, Russel  
APPLICANT: Yoneda, Toshiyuki  
TITLE OF INVENTION: Integrin Binding Motif Containing  
Peptides and Methods of Treating Skeletal Diseases  
FILE REFERENCE: BEAR-006CIP  
CURRENT APPLICATION NUMBER: US/09/812,485A  
CURRENT FILING DATE: 2001-03-19  
PRIOR APPLICATION NUMBER: 09/641,034  
PRIOR FILING DATE: 2000-08-16  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 32  
LENGTH: 40  
TYPE: PRF  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: peptidic compound  
US-09-812-485A-32

Query Match 45.2%; Score 57; DB 9; Length 40;  
Best Local Similarity 100.0%; Pred. No. 0.023;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 FSGDQPFKD 23  
 Db 1 FSGDQPFKD 10

RESULT 15

US-09-812-485A-27  
 ; Sequence 27, Application US/09812485A  
 ; Publication No. US20020197267A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kumagai, Yoshiaki  
 ; APPLICANT: Blacher, Russel  
 ; APPLICANT: Yoneda, Yoshiyuki  
 ; TITLE OF INVENTION: Integrin Binding Motif Containing  
 ; TITLE OF INVENTION: Peptides and Methods of Treating Skeletal Diseases  
 ; FILE REFERENCE: BEAR-006CIP  
 ; CURRENT APPLICATION NUMBER: US/09/812,485A  
 ; CURRENT FILING DATE: 2001-03-19  
 ; PRIOR APPLICATION NUMBER: 09/641,034  
 ; PRIOR FILING DATE: 2000-08-16  
 ; NUMBER OF SEQ ID NOS: 50  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 27  
 ; LENGTH: 45  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: peptidic compound  
 US-09-812-485A-27

Query Match 45.2% Score 57; DB 9; Length 45;  
 Best Local Similarity 100.0%; Pred. No. 0.027;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 FSGDQPFKD 23  
 Db 1 FSGDQPFKD 10

Search completed: January 17, 2003, 13:22:03  
 Job time : 10 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 13:17:52 ; Search time 15 Seconds

(without alignments)  
112,414 Million cell updates/sec

Title: US-09-641-034-47

Perfect score: 126

Sequence: 1 TDQERGDNDISPFSGDGPFPKD 23

Scoring table:

BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 273288 seqs, 7331625 residues

Total number of hits satisfying chosen parameters: 273288

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New:\*

- 1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*
- 7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	37.3	815	6	US-10-092-411A-3614
2	46.5	36.9	1141	6	US-10-145-087A-437
3	46.5	36.9	1141	6	US-10-143-031A-437
4	46.5	36.9	1141	6	US-10-145-092A-437
5	46.5	36.9	1141	6	US-10-162-522A-437
6	46.5	36.9	1141	6	US-10-165-038A-437
7	46.5	36.9	1141	6	US-10-165-353-437
8	46.5	36.9	1141	6	US-10-170-481A-437
9	46.5	36.9	1141	6	US-10-172-039A-437
10	46.5	36.9	1141	6	US-10-145-016A-437
11	46.5	36.9	1141	6	US-10-145-088A-437
12	46.5	36.9	1141	6	US-10-145-129A-437
13	46.5	36.9	1141	6	US-10-125-823A-120
14	46.5	36.9	1141	6	US-10-165-353A-437
15	46.5	36.9	1141	6	US-10-205-892-120
16	46.5	36.9	1141	6	US-10-174-575-120
17	46.5	36.9	1141	6	US-10-174-575A-120
18	46.5	36.9	1141	6	US-10-162-521A-437
19	46.5	36.9	1141	6	US-10-167-600-437
20	46.5	36.9	1141	6	US-10-187-755-120
21	46.5	36.9	1141	6	US-10-013-924A-437
22	46.5	36.9	1141	6	US-10-013-928A-437
23	46.5	36.9	1141	6	US-10-167-600A-437
24	46.5	36.9	1141	6	US-10-187-749-120
25	46.5	36.9	1141	6	US-10-187-749-120
26	46.5	36.9	1165	5	US-09-724-676-75168

27	46.5	36.9	1165	5	US-09-724-676A-75168	Sequence 75168, A
28	46.5	36.9	1172	5	US-09-724-676-75166	Sequence 75166, A
29	46.5	36.9	1172	5	US-09-724-676A-75166	Sequence 75166, A
30	46.5	36.9	1200	5	US-09-724-676-75170	Sequence 75170, A
31	46.5	36.9	1200	5	US-09-724-676A-75170	Sequence 75170, A
32	46.5	36.9	1203	5	US-09-724-676-75169	Sequence 75169, A
33	46.5	36.9	1203	5	US-09-724-676A-75169	Sequence 75169, A
34	46.5	36.9	1210	5	US-09-724-676-75167	Sequence 75167, A
35	46.5	36.9	1210	5	US-09-724-676A-75167	Sequence 75167, A
36	46.5	36.9	1238	5	US-09-724-676-75171	Sequence 75171, A
37	46.5	36.9	1238	5	US-09-724-676A-75171	Sequence 75171, A
38	46	36.5	648	1	PCT-US02-32727-15272	Sequence 15272, A
39	46	36.5	648	6	US-10-057-498-15272	Sequence 15272, A
40	43	34.1	2476	5	US-09-724-676-91032	Sequence 91032, A
41	43	34.1	2476	5	US-09-724-676A-91032	Sequence 91032, A
42	43	34.1	2548	5	US-09-724-676-91030	Sequence 91030, A
43	43	34.1	2548	5	US-09-724-676A-91030	Sequence 91030, A
44	43	34.1	2826	5	US-09-724-676-91028	Sequence 91028, A
45	43	34.1	2826	5	US-09-724-676A-91028	Sequence 91028, A

## ALIGNMENTS

RESULT 1  
US-10-092-411A-3614  
Sequence 3614, Application US/10092411A  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 032796-101  
CURRENT APPLICATION NUMBER: US/10/092.411A  
CURRENT FILING DATE: 2002-03-07  
PRIOR APPLICATION NUMBER: US 09/134,001  
PRIOR FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5676  
SEQ ID NO 3614  
LENGTH: 815  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-10-092-411A-3614

Query Match 37.3%; Score 47; DB 6; Length 815;  
Best Local Similarity 58.8%; Pred. No. 78;  
Matches 10; Conservative 7; Indels 0; Gaps 0;

QY 7 GDNDISPFSGDGPFPKD 23  
DB 481 GDEDLPVSGDGPFPKD 497

## RESULT 2

US-10-145-087A-437  
Sequence 437, Application US/10145087A  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Baker Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Ealon, Dan  
APPLICANT: Ferrara, Napoleon  
APPLICANT: Fliviaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gottfredsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.

```

: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kijavlin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James.
: APPLICANT: Paonl, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630PIC47
: CURRENT APPLICATION NUMBER: US/10/145,087A
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/074450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 624
: SEQ ID NO 437
: LENGTH: 1141
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-145-087A-437

Query Match          36.9%; Score 46.5; DB 6; Length 1141;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

QY      2 DLQERGDNDI---SPFGSGGQPF 21
      11 : 1 11 : 11 111 : 1
DB      394 DLNQGFPDIAGVAPFDCGKVF 416

RESULT 3
US-10-143-031A-437
: Sequence 437, Application US/10143031A
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
```

```

: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kijavlin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James.
: APPLICANT: Paonl, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630PIC39
: CURRENT APPLICATION NUMBER: US/10/143,031A
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2002-10-10
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/074450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 624
: SEQ ID NO 437
: LENGTH: 1141
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-143-031A-437

Query Match          36.9%; Score 46.5; DB 6; Length 1141;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

QY      2 DLQERGDNDI---SPFGSGGQPF 21
      11 : 1 11 : 11 111 : 1
DB      394 DLNQGFPDIAGVAPFDCGKVF 416

RESULT 4
US-10-145-092A-437
: Sequence 437, Application US/10145092A
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
```





```

: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kljavin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James.
: APPLICANT: Paonli, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630PIC29
: CURRENT APPLICATION NUMBER: US/10/165.038A
: CURRENT FILING DATE: 2002-10-10
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 624
: SEQ ID NO 437
: LENGTH: 1141
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-165-038A-437

Query Match          36.9% Score 46.5; DB 6; Length 1141;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

OY 2 DLOERGNDI---SPFSGDQPF 21
||:| |||: ||| |||: |
Db 394 DLNODGFPDIAGAFPDGDKYF 416

RESULT 7
US-10-165-353-437
: Sequence 437, Application US/10165353
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerltsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
```

```

: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kljavin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James.
: APPLICANT: Paonli, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630PIC40
: CURRENT APPLICATION NUMBER: US/10/165.353
: CURRENT FILING DATE: 2002-10-10
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 624
: SEQ ID NO 437
: LENGTH: 1141
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-165-353-437

Query Match          36.9% Score 46.5; DB 6; Length 1141;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

OY 2 DLOERGNDI---SPFSGDQPF 21
||:| |||: ||| |||: |
Db 394 DLNODGFPDIAGAFPDGDKYF 416

RESULT 8
US-10-170-481A-437
: Sequence 437, Application US/10170481A
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerltsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
```

```

: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kljavin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James.
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630P1C53
: CURRENT APPLICATION NUMBER: US/10/170,481A
: CURRENT FILING DATE: 2002-10-10
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 624
: SEQ ID NO 437
: LENGTH: 1141
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-170-481A-437

Query Match          36.9% Score 46.5; DB 6; Length 1141;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

QY      2 DLQERGDNDI---SPFGSDGQPF 21
      11 : 1 11 : 11 111 : 1
Db      394 DLNDGFPDIAVGAFPDGDGRVF 416

RESULT 9
US-10-172-039A-437
: Sequence 437, Application US/10172039A
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
```

```

: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kljavin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James.
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630P1C30
: CURRENT APPLICATION NUMBER: US/10/172,039A
: CURRENT FILING DATE: 2002-10-10
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 624
: SEQ ID NO 437
: LENGTH: 1141
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-172-039A-437

Query Match          36.9% Score 46.5; DB 6; Length 1141;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

QY      2 DLQERGDNDI---SPFGSDGQPF 21
      11 : 1 11 : 11 111 : 1
Db      394 DLNDGFPDIAVGAFPDGDGRVF 416

RESULT 10
US-10-145-016A-437
: Sequence 437, Application US/10145016A
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
```

```

: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kljavin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James.
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630P1C52
: CURRENT FILING DATE: US/10/145.016A
: PRIOR FILING DATE: 2001-10-18
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 624
: SEQ ID NO 437
: LENGTH: 1141
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-10-145-016A-437

Query Match          36.9%; Score 46.5; DB 6; Length 1141;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

OY      2 DLORGDNDI---SPFGSDGQPF 21
Db      394 DLNODGFPIAVGAPFDGDKVF 416

RESULT 11
: US-10-145-088A-437
: Sequence 437, Application US/10145088A
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gertlisen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.

```

```

: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kljavin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James.
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630P1C49
: CURRENT FILING DATE: US/10/145.088A
: PRIOR FILING DATE: 2002-10-10
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 624
: SEQ ID NO 437
: LENGTH: 1141
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-10-145-088A-437

Query Match          36.9%; Score 46.5; DB 6; Length 1141;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

OY      2 DLORGDNDI---SPFGSDGQPF 21
Db      394 DLNODGFPIAVGAPFDGDKVF 416

RESULT 12
: US-10-145-129A-437
: Sequence 437, Application US/10145129A
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gertlisen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.

```

```

: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kijavlin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James;
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630PIC51
: CURRENT APPLICATION NUMBER: US/10/145,129A
: CURRENT FILING DATE: 2002-10-10
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 624
: SEQ ID NO 437
: LENGTH: 1141
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-10-145-129A-437

Query Match      36.9%; Score 46.5; DB 6; Length 1141;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

QY      2 DLQERGDNDI---SPFSGDCQPF 21
      11 : 1 11 : 11 111 : 1
Db      394 DLNODGPDIVAGAPFDGDKVF 416

RESULT 13
: US-10-125-923A-120
: Sequence 120, Application US/10125923A
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Chen, Jian
: APPLICANT: Desnoyers, Luc
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Smith, Victoria
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: TITLE OF INVENTION: ACIDS ENCODING THE SAME

```

```

: FILE REFERENCE: P3430RIC79
: CURRENT APPLICATION NUMBER: US/10/125,923A
: CURRENT FILING DATE: 2002-01-15
: PRIOR APPLICATION NUMBER: 10/052586
: PRIOR FILING DATE: 2002-01-15
: PRIOR APPLICATION NUMBER: 60/059263
: PRIOR FILING DATE: 1997-09-18
: PRIOR APPLICATION NUMBER: 60/059266
: PRIOR FILING DATE: 1997-09-18
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/063120
: PRIOR FILING DATE: 1997-10-24
: PRIOR APPLICATION NUMBER: 60/063121
: PRIOR FILING DATE: 1997-10-24
: PRIOR APPLICATION NUMBER: 60/063486
: PRIOR FILING DATE: 1997-10-21
: PRIOR APPLICATION NUMBER: 60/063540
: PRIOR FILING DATE: 1997-10-28
: PRIOR APPLICATION NUMBER: 60/063541
: PRIOR FILING DATE: 1997-10-28
: PRIOR APPLICATION NUMBER: 60/063544
: PRIOR FILING DATE: 1997-10-28
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 612
: SEQ ID NO 120
: LENGTH: 1141
: TYPE: PRT
: ORGANISM: Homo Sapien
: US-10-125-923A-120

Query Match      36.9%; Score 46.5; DB 6; Length 1141;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

QY      2 DLQERGDNDI---SPFSGDCQPF 21
      11 : 1 11 : 11 111 : 1
Db      394 DLNODGPDIVAGAPFDGDKVF 416

RESULT 14
: US-10-165-353A-437
: Sequence 437, Application US/10165353A
: GENERAL INFORMATION:
: APPLICANT: Ashtkenazi, AVI
: APPLICANT: Baker, Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Geo, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kijavlin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James;
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: TITLE OF INVENTION: ACIDS ENCODING THE SAME

```

```

; FILE REFERENCE: P2630P1C40
; CURRENT APPLICATION NUMBER: US/10/165.353A
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 624
; SEQ ID NO 437
; LENGTH: 1141
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-165-353A-437

```

```

Query Match          36.9%; Score 46.5; DB 6; Length 1141;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

```

```

OY 2 DLQERGDNDI---SPFSGDGP 21
Db 394 DLNDGFPDIAVGAPFDGDKVF 416

```

```

RESULT 15
US-10-205-892-120
; Sequence 120, Application US/10205892
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C517
; CURRENT APPLICATION NUMBER: US/10/205.892
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486

```

```

; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 120
; LENGTH: 1141
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-205-892-120

```

```

Query Match          36.9%; Score 46.5; DB 6; Length 1141;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

```

```

OY 2 DLQERGDNDI---SPFSGDGP 21
Db 394 DLNDGFPDIAVGAPFDGDKVF 416

```

```

Search completed: January 17, 2003, 13:21:47
Job time : 17 secs

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 13:16:38 ; Search time 16 Seconds

(without alignments)  
138.193 Million cell updates/sec

Title: US-09-641-034-47

Perfect score: 126

Sequence: 1 TDLOERGNDNDISPPSGDGPFRKD 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	41.3	301	2	S16070
2	52	41.3	320	2	JC5727
3	49.5	39.3	1259	2	S36126
4	48	38.1	277	2	AD3322
5	48	38.1	385	2	D72858
6	48	38.1	385	2	T41811
7	47	37.3	444	2	B83891
8	47	37.3	444	2	AH0410
9	47	37.3	825	2	B89944
10	46.5	36.9	1106	2	S38783
11	46.5	36.9	1135	2	161186
12	46.5	36.9	1137	2	JC5950
13	46	36.5	308	2	AD3154
14	46	36.5	365	2	E98133
15	45.5	36.1	347	2	A83170
16	45.5	36.1	1260	1	S05479
17	45	35.7	158	2	A11365
18	45	35.7	357	2	C84856
19	45	35.7	372	2	S59499
20	45	35.7	473	2	J00802
21	45	35.7	913	2	T38173
22	45	35.7	933	2	D82885
23	44.5	35.3	698	2	D65210
24	44.5	35.3	698	2	D91255
25	44.5	35.3	698	2	H86095
26	44	34.9	348	2	A49876
27	44	34.9	475	1	S15337
28	44	34.9	475	2	AC0528
29	44	34.9	683	2	D83511

30	43.5	34.5	284	2	C71507	probable L2 riboso
31	43.5	34.5	284	2	A81661	ribosomal protein
32	43.5	34.5	631	2	D38162	cof protein - pse
33	43.5	34.5	1053	2	S44250	integrin alpha-5 c
34	43.5	34.5	1115	2	T09403	integrin alpha cha
35	43.5	34.5	1115	2	T09433	integrin alpha cha
36	43	34.1	140	2	G84608	En/Spm-like transp
37	43	34.1	300	2	F90274	hypothetical prote
38	43	34.1	335	2	G72258	conserved hypotnet
39	43	34.1	629	2	T39285	probable transmemb
40	43	34.1	684	1	RNA6VS	transcription init
41	43	34.1	684	2	G97619	RNA polymerase sig
42	43	34.1	684	2	AF2842	RNA polymerase sig
43	43	34.1	762	2	E87592	hypothetical prote
44	43	34.1	779	2	G87573	xanthine dehydroge
45	43	34.1	891	2	E96590	hypothetical prote

## ALIGNMENTS

RESULT 1  
S16070  
porin - Rhodobacter capsulatus  
C:Species: Rhodobacter capsulatus  
C:Date: 21-Nov-1993 #sequence\_revision 21-Jul-1995 #text\_change 21-Jul-1995  
R:Schultz, E.; Kreusch, A.; Nestel, U.; Schulz, G.E.  
Eur. J. Biochem. 199, 587-594, 1991  
A:Title: Primary structure of porin from Rhodobacter capsulatus.  
A:Reference number: S16070; MUID:91330909; PMID:1651239  
A:Accession: S16070  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-301 <EUF>  
A:Experimental source: strain 37b4  
A>Note: the source is designated as Rhodobacter capsulatus

Query Match 41.3% Score 52; DB 2; Length 301;  
Best Local Similarity 58.8% Pred. No. 3.1;  
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
Oy 1 TDLOERGNDISPPSGD 17  
Db 92 TDLDGRGNDIPYLTGD 108

RESULT 2  
JC5727  
porin precursor - Rhodobacter capsulatus  
C:Species: Rhodobacter capsulatus  
C:Date: 09-Dec-1997 #sequence\_revision 23-Jan-1998 #text\_change 07-May-1999  
C:Accession: JC5727; PC4414  
R:Trieschmann, M.D.A.; Patus, F.; Tadros, M.H.  
Gene 183, 61-68, 1996  
A:Title: Molecular characterization and organization of porin from Rhodobacter capsu.

A:Reference number: JC5727; MUID:97149280; PMID:8996088  
A:Accession: JC5727  
A:Molecule type: DNA  
A:Residues: 1-320 <TRI>  
A:Cross-references: GB:U57653  
A:Experimental source: strain 37B4  
A:Accession: PC4414  
A:Molecule type: protein  
A:Residues: 67-89;131-158;191-218;252-277 <RP2>  
C:Comment: This protein is the mutant of porin from Rhodobacter capsulatus strain 37I  
ion, which are located exclusively on transmembrane strands.  
C:Genetics:

A:Gene: porCa  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-320/Product: porin #status predicted <MNT>

Query Match 41.3% Score 52; DB 2; Length 320;

Best Local Similarity 58.8%; Pred. No. 3.3;  
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
OY 1 TDLOERGNDISPFGSD 17  
|||:|||||:|||||  
DB 112 TDLOERGNDISPFGSD 128

## RESULT 3

S36126

neural cell adhesion molecule L1 - rat

M:Alternate names: nerve growth factor-inducible large external glycoprotein; NILE glyco

C:Species: Rattus norvegicus (Norway rat)

C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 20-Aug-1999

R:Accession: S36126; S17655; A60917; A30326

R:Miura, M.; Kobayashi, M.; Asou, H.; Umemura, K.

A:Title: Molecular cloning of cDNA encoding the rat neural cell adhesion molecule L1. T

A:Reference number: S17655; MUID:91372414; PMID:1894011

A:Accession: S36126

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1259 &lt;MID&gt;

A:Cross-references: EMBL:X59149

A:Accession: S17655

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1178, 1183-1259 &lt;MID&gt;

A:Cross-references: EMBL:X59149; NID:956740; PIDN:CMA41860.1; PID:956741

R:Prince, J.T.; Milona, N.; Stallcup, W.B.

J. Neurosci. 9, 1825-1834, 1989

A:Title: Characterization of a partial cDNA clone for the NILE glycoprotein and identifi

A:Reference number: A60917; MUID:89257627; PMID:2723751

A:Accession: A60917

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1159-1199, 'G', 1201-1235, 'K', 1237 &lt;PRI&gt;

A:Note: this paper appeared earlier, with printing errors, as reference A30326

R:Prince, J.T.; Milona, N.; Stallcup, W.B.

J. Neurosci. 9, 876-883, 1989

A:Title: Characterization of a partial cDNA clone for the NILE glycoprotein and identifi

A:Reference number: A30326; MUID:89177485; PMID:2466966

A:Contents: annotation

A:Note: this paper was reprinted as reference A60917 to correct the omission of several

C:Comment: This sequence of this surface-accessible glycoprotein differs at only two pos

C:Superfamily: neural cell adhesion molecule L1; fibronectin type III repeat homology; 1

C:Keywords: cell adhesion; duplication; glycoprotein; membrane protein

F:531-592/Domains: immunoglobulin homology &lt;IMM&gt;

## Query Match

Best Local Similarity 39.3%; Score 49.5; DB 2; Length 1259;

Matches 12; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

## OY 2

DLOERGNDISPFGSDQ 19

|||||||:|:|

DB 558 DLOERGNDISPFGSDQ 574

## RESULT 4

AD3322

lysosome (BC 3.2.1.17) - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 19-Apr-2002

C:Accession: AD3322

R:DelVecchio, V.G.; Kapural, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova,

; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Leless

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A:Reference number: AD3322; PMID:11756688

A:Accession: AD3322

A:Status: preliminary

A:Molecule type: DNA

A:Experimental source: isolate T3

A:Residues: 1-277 <KUR>  
A:Cross-references: GB:A6008917; PIDN:AAL51743.1; PID:g17982481; GSPDB:GN00190  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BME10562  
A:Map position: 1  
C:Keywords: glycosidase; hydrolase

Query Match 38.1%; Score 48; DB 2; Length 277;  
Best Local Similarity 52.9%; Pred. No. 11;

Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

## OY 6

RCDNISPFGSGOPFK 22

|||||:|:|

DB 252 RCDNISPFGSGOPFK 268

## RESULT 5

D72858

late expression factor 3 - Autographa californica nuclear polyhedrosis virus

C:Species: Autographa californica nuclear polyhedrosis virus, AcMNPV

A:Note: dsDNA virus

C:Date: 12-Nov-1999 #sequence\_revision 12-Nov-1999 #text\_change 24-Nov-1999

R:Accession: D72858; A40677

R:Avres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.

A:Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis vir

A:Reference number: A72850; MUID:94303173; PMID:8030224

A:Accession: D72858

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-385 &lt;AVR&gt;

A:Cross-references: GB:L22858; NID:9510708; PIDN:AAA66697.1; PID:9559136

R:Li, Y.; Passarelli, A.L.; Miller, L.K.

J. Virol. 67, 5260-5268, 1993

A:Title: Identification, sequence, and transcriptional mapping of Ief-3, a baculoviru

A:Reference number: A40677; MUID:93353600; PMID:8350397

A:Accession: A40677

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-385 &lt;LIA&gt;

A:Cross-references: GB:L18873; NID:9349019; PIDN:AAA02964.1; PID:9349020

A:Note: authors translated the codon ACA for residue 92 as Tyr

C:Genetics:

A:Gene: Ac-Ief3

C:Keywords: transcription regulation

## Query Match

Best Local Similarity 38.1%; Score 48; DB 2; Length 385;

Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

## OY 8

DNDISPFGSDGQPKD 23

|||||:|:|

DB 61 DNDISPFGSDGQPKD 76

## RESULT 6

T41811

LEF-3 orf67 - Bombyx mori nuclear polyhedrosis virus (isolate T3)

C:Species: Bombyx mori nuclear polyhedrosis virus, BmNPV

A:Variety: isolate T3

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jul-2000

C:Accession: T41811

R:Gomi, S.; Maizumi, K.; Maeda, S.

J. Gen. Virol. 80, 1323-1337, 1999

A:Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.

A:Reference number: 222020; MUID:99281911; PMID:10355780

A:Accession: T41811

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-385 &lt;KAM&gt;

A:Cross-references: EMBL:L33180; NID:93745835; PIDN:AMC63740.1; PID:93745893

A:Experimental source: isolate T3

C:Genetics:  
A:Note: lrf-3

Query Match 38.1%; Score 48; DB 2; Length 385;  
Best Local Similarity 56.2%; Pred. No. 17;  
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 8 DNDISPFSGDGPFRD 23  
DB 61 DNKIOEYXGDSQSFKD 76

RESULT 7  
B83891  
Intracellular alkaline serine proteinase aprX [Imported] - *Bacillus halodurans* (strain C)  
C:Species: *Bacillus halodurans*  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: B83891  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: B83891  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-444 <STO>  
A:Cross-references: GB:BA000004; NID:g10174345; PIDN:BA05649.1; GSPDB:GN00  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: aprX

Query Match 37.3%; Score 47; DB 2; Length 444;  
Best Local Similarity 47.1%; Pred. No. 28;  
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 2 DLOERGDNDISFGDG 18  
DB 319 DTTREDDDVAFSSRG 335

RESULT 8  
AH0410  
Global stress requirement protein Gsra [Imported] - *Yersinia pestis* (strain CO92)  
C:Species: *Yersinia pestis*  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 17-May-2002  
C:Accession: AH0410  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,  
Nature 413, 533-537, 2001  
A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AH0410  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-481 <KUR>  
A:Cross-references: GB:AL590842; PIDN:CAC92612.1; PID:g15981308; GSPDB:GN00175  
C:Genetics:  
A:Gene: gsra  
C:Superfamily: Helicobacter serine proteinase

Query Match 37.3%; Score 47; DB 2; Length 481;  
Best Local Similarity 62.5%; Pred. No. 30;  
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

OY 7 GDNDISPFGDGPFRK 22  
DB 80 GDN--SPFGDGSFPQ 93

RESULT 9  
B89944  
hypothetical protein SA1447 [Imported] - *Staphylococcus aureus* (strain N315)

C:Species: *Staphylococcus aureus*  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C:Accession: B89944  
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.;  
ma, A.; Mizutani-H, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu,  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsuki, K.  
Lancet 357, 1225-1240, 2001  
A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.  
A:Reference number: A89758; MUID:21311952; PMID:1141816  
A:Accession: B89944  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-825 <KUR>  
A:Cross-references: GB:BA000018; PID:g13701417; PIDN:BA042711.1; GSPDB:GN00149  
A:Experimental source: strain N315  
C:Genetics:  
A:Gene: SA1447

Query Match 37.3%; Score 47; DB 2; Length 825;  
Best Local Similarity 58.8%; Pred. No. 55;  
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 7 GDNDISPFGDGPFRD 23  
DB 476 GDEDDLPVSGQVFRD 492

RESULT 10  
S38783  
Integrin alpha chain - rat (fragment)  
C:Species: *Rattus norvegicus* (Norway rat)  
C:Date: 22-Jan-1994 #sequence\_revision 14-Jul-1995 #text\_change 29-Sep-1999  
C:Accession: S38783; S23600  
R:Kaufman, S.J.  
submitted to the EMBL Data Library, March 1992  
A:Reference number: S38783  
A:Accession: S38783  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1106 <KAU>  
A:Cross-references: EMBL:X65036; NID:g56392; PIDN:CAA6170.1; PID:g56393  
J:Song, W.K.; Wang, W.; Foster, R.F.; Bjelseth, D.A.; Kaufman, S.J.  
Cell Biol. 117, 643-657, 1992  
A:Title: H36-alpha7 is a novel integrin alpha chain that is developmentally regulated  
A:Reference number: S23600; MUID:92242309; PMID:1315319  
A:Accession: S23600  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-116, 'R', 118-349, 'D', 351-804, 'R', 806, 'V', 808-1106 <SON>  
A:Cross-references: EMBL:X65036  
C:Superfamily: Integrin alpha-2b chain

Query Match 36.9%; Score 46.5; DB 2; Length 1106;  
Best Local Similarity 47.8%; Pred. No. 91;  
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

OY 2 DLOERGDNDI---SPFGDGPFR 21  
DB 362 DLNDGPPDIAGAFPDGKRVF 384

RESULT 11  
I61186  
alpha-7 integrin - mouse  
C:Species: *Mus musculus* (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 29-Sep-1999  
C:Accession: I61186  
J:Zlober, B.L.; Vu, M.P.; Maleh, N.; Crawford, J.; Lin, C.S.; Krammer, R.H.  
J. Biol. Chem. 268, 26773-26783, 1993  
A:Title: Alternative extracellular and cytoplasmic domains of the integrin alpha 7 su  
A:Reference number: A49691; MUID:94075378; PMID:8253814  
A:Accession: I61186  
A:Status: preliminary; translated from GR/EMBL/DBJ





Fri Jan 17. 13:11:06 2003

us-09-641-034-47.rpr

Page 5

```

A:Reference number: AB2950; MWID:20437337; PMID:10984043
A:Accession: AB3170
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-347 <5'3'>
A:Cross-references: GB:AE004799; GB:AE004091; NID:99949981; PIDN:AMG07211.1; GSPDB:GN001
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: queA; PA3824
C:Superfamily: Escherichia coli tRNA ribosyltransferase-isomerase
C:Keywords: Intramolecular transferase; isomerase

Query Match          36.1%; Score 45.5; DB 2; Length 347;
Best Local Similarity 42.9%; Pred. No. 36;
Matches 9; Conservative 3; Mismatches 2; Indels 7; Gaps 1;

Oy      8 DNDISPFSGD-----GQPF 21
         |::|:|:|:|:|:|:|:|:|
Db      263 DGELEKPFSGDIDFIYRGRPF 283

```

Search completed: January 17, 2003, 13:17:03  
Job time : 18 secs

Fri Jan 17 13:11:06 2003

us-09-641-034-47.rsp

Page 1

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 13:16:38 ; Search time 10 Seconds

(without alignments)  
95.396 Million cell updates/sec

Title: US-09-641-034-47

Sequence: 1 TDQERGDNDISPFSGDGPFKD 23

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	41.3	301	1	P31243 rhodobacter
2	49.5	39.3	1259	1	Q05695 rattus norv
3	48	38.1	385	1	P41453 autographa
4	46.5	36.9	1106	1	O61738 mus musculu
5	46.5	36.9	1179	1	O61738 mus musculu
6	46.5	36.9	1181	1	O61738 mus musculu
7	45.5	36.1	1015	1	O61738 mus musculu
8	45.5	36.1	1015	1	O61738 mus musculu
9	45.5	36.1	1015	1	O61738 mus musculu
10	45.5	36.1	1015	1	O61738 mus musculu
11	45	35.7	309	1	P31627 mus musculu
12	45	35.7	309	1	P31627 mus musculu
13	45	35.7	309	1	P31627 mus musculu
14	45	35.7	309	1	P31627 mus musculu
15	45	35.7	309	1	P31627 mus musculu
16	44.5	35.3	688	1	Q10366 schistosom
17	44.5	35.3	1050	1	O06274 xenopus lae
18	44	34.9	475	1	O06274 xenopus lae
19	44	34.9	475	1	O06274 xenopus lae
20	43.5	34.5	284	1	Q39317 chlamydia m
21	43.5	34.5	284	1	Q39317 chlamydia m
22	43.5	34.5	284	1	Q39317 chlamydia m
23	43.5	34.5	284	1	Q39317 chlamydia m
24	43.5	34.5	284	1	Q39317 chlamydia m
25	43	34.1	684	1	P33452 agrobacteri
26	43	34.1	684	1	P33452 agrobacteri
27	43	34.1	684	1	P33452 agrobacteri
28	43	34.1	684	1	P33452 agrobacteri
29	42.5	33.7	1000	1	ITAS_DROME
30	42.5	33.7	1000	1	ITAS_DROME
31	42	33.3	119	1	GVOI_HALNI
32	42	33.3	120	1	GVOI_HALNI
33	42	33.3	224	1	DCL_LYCES

#### ALIGNMENTS

34	42	33.3	269	1	MHPD_ECOLI	P77608 escherichia
35	42	33.3	359	1	HRCA_RHIME	O925K1 rhizobium m
36	42	33.3	445	1	SLP2_DROME	P31031 drosophila
37	42	33.3	557	1	FLIE_RHIME	O54239 rhizobium m
38	42	33.3	823	1	SYL_DEIRA	O958FO delinococcus
39	42	33.3	1192	1	METH_MYCTU	O33259 mycobacteri
40	42	33.3	1756	1	TRIL_ECOLI	P14565 escherichia
41	41.5	32.9	232	1	VHED_BPT3	P20313 bacterioph
42	41.5	32.9	265	1	FAEH_ECOLI	P33782 escherichia
43	41.5	32.9	318	1	RSK_SCHPO	O60116 schistosom
44	41.5	32.9	811	1	ENAN_BPKIE	P49714 bacterioph
45	41.5	32.9	1025	1	ITAB_HUMAN	P53708 homo sapien

RESULT 1  
P31243 rhodobacter  
ID AC P31243: STANDARD: PRT: 301 AA.  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 01-OCT-1994 (Rel. 30, Last annotation update)  
DE Porin.  
OS Rhodobacter capsulatus (Rhodopseudomonas capsulatus).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;  
CC Rhodobacter.  
OX NCBI\_TaxID=1061;  
RN [1]  
RP SEQUENCE.  
RC STRAIN-DSM 938 / 37B4;  
RX MEDLINE=9130909; PubMed=1651239;  
RA Schlicht E., Kreusch A., Nestel U., Schulz G.E.;  
RT "Primary structure of porin from Rhodobacter capsulatus.";  
RL Eur. J. Biochem. 199;587-594(1991).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).  
RC STRAIN-DSM 938 / 37B4;  
RX MEDLINE=9036791; PubMed=2165921;  
RA Weiss M.S., Wacker T., Weckesser J., Walte W., Schulz G.E.;  
RT "The three-dimensional structure of porin from Rhodobacter capsulatus at 3-A resolution.";  
RL FEBS Lett. 267:268-272(1990).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).  
RC STRAIN-DSM 938 / 37B4;  
RX MEDLINE=91192174; PubMed=1707373;  
RA Weiss M.S., Kreusch A., Schlicht E., Nestel U., Walte W., Weckesser J., Schulz G.E.;  
RT "The structure of porin from Rhodobacter capsulatus at 1.8-A resolution.";  
RL FEBS Lett. 280:379-382(1991).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).  
RC MEDLINE=93021091; PubMed=1328651;  
RA Weiss M.S., Schulz G.E.;  
RT "Structure of porin refined at 1.8-A resolution.";  
RL J. Mol. Biol. 227:493-509(1992).  
CC -1- FUNCTION: FORMS CHANNELS THAT ALLOW THE PASSIVE DIFFUSION OF SMALL HYDROPHILIC SOLUTIONS UP TO AN EXCLUSION LIMIT OF ABOUT 0.6 kDa.  
CC -1- SUBUNIT: HOMOTRIMER.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane. PEATED SHEET BARREL (171 RESIDUES), THREE SHORT ALPHA-HELICES (18 RESIDUES) AND 13 HYDROGEN-BONDED REVERSE TURNS (26 RESIDUES).  
CC PIR: S16070; S16070.  
DR PDB: 2POR; 15-JUL-93.  
DR PDB: 3POR; 15-JUL-93.  
KW Outer membrane; Transmembrane; Porin; 3D-structure.  
FT STRAND 2 14  
FT STRAND 19 34  
FT TURN 36 37

[illegible]

FT CARBOHYD 1106 1106 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARSPLIC 1179 1182 MISSING (IN ISOFORM 2).  
 SQ SEQUENCE 1259 AA: 140934 MW: 0F12A7C4415F3C08 CRC64;  
 Query Match 39.3%; Score 49.5; DB 1; Length 1259;  
 Best Local Similarity 66.7%; Pred. No. 13;  
 Matches 12; Conservative 1; Mismatches 4; Indels 1; Gaps 1;  
 QY 2 DLQERGDNDISFGSDGQ 19  
 DB 558 DLQERGDSD-KYFIEDGQ 574  
 RESULT 3  
 LEF3\_NPVAC STANDARD: PRT: 385 AA.  
 ID LEF3\_NPVAC  
 AC P41453;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Late expression factor 3.  
 GN LEF-3.  
 OS Autographa californica nuclear polyhedrosis virus (AcNPV).  
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
 CC Nucleopolyhedrovirus.  
 CX NCBI\_TaxID=46015;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-L1;  
 RX MEDLINE=93353600; PubMed=8350397;  
 RA Li Y., Passarelli A.L., Miller L.K.;  
 RT "Identification, sequence, and transcriptional mapping of lef-3, a  
 RT baculovirus gene involved in late and very late gene expression.";  
 RL J. Virol. 67:5260-5268(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-C6;  
 RX MEDLINE=94303173; PubMed=8030224;  
 RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;  
 RT "The complete DNA sequence of Autographa californica nuclear  
 RT polyhedrosis virus.";  
 RL Virology 202:586-605(1994).  
 CC -1- FUNCTION: REQUIRED FOR LATE AND VERY LATE GENE EXPRESSION. LEF-3  
 CC COULD BE A SINGLE STRANDED DNA-BINDING PROTEIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L18873; AAA02964.1;  
 DR EMBL: L22858; AAA66697.1;  
 DR PIR: A40677; A40677.  
 KW Early protein; transcription regulation; DNA-binding.  
 SO SEQUENCE 385 AA: 44551 MW: 9A25ECD7BA7EBDF1 CRC64;  
 Query Match 38.1%; Score 48; DB 1; Length 385;  
 Best Local Similarity 56.2%; Pred. No. 5.5;  
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
 QY 8 DNDISFGSDGQPFKD 23  
 DB 61 DNKIQEYVYGDGSGFKD 76  
 RESULT 4  
 ITA7\_RAT STANDARD: PRT: 1106 AA.  
 ID ITA7\_RAT  
 AC 063258; 063026; 063027;  
 DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Integrin alpha-7 (H36-alpha7).  
 GN ITGA7.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 CX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM ALPHA-7X1B).  
 RC TISSUE-Skeletal muscle;  
 RX MEDLINE=92242309; PubMed=1315319;  
 RA Song W.K., Wang W., Foster R.F., Bleiser D.A., Kaufman S.J.;  
 RT "H36-alpha 7 is a novel integrin alpha chain that is developmentally  
 RT regulated during skeletal myogenesis.";  
 RL J. Cell Biol. 117:643-657(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS ALPHA-7X1A AND ALPHA-7X1C).  
 RC TISSUE-Skeletal muscle;  
 RX MEDLINE=94171924; PubMed=8126096;  
 RA Song W.K., Wang W., Sato H., Bleiser D.A., Kaufman S.J.;  
 RT "Expression of alpha 7 integrin cytoplasmic domains during skeletal  
 RT muscle development: alternate forms, conformational change, and  
 RT homologues with serine/threonine kinases and tyrosine phosphatases.";  
 RL J. Cell Sci. 106:1139-1152(1993).  
 RN [3]  
 RP TISSUE SPECIFICITY.  
 RX MEDLINE=96197133; PubMed=8626012;  
 RA Martin P.T., Kaufman S.J., Kramer R.H., Sanes J.R.;  
 RT "Synaptic integrins in developing adult, and mutant muscle: selective  
 RT association of alpha1, alpha7A, and alpha7B integrins with the  
 RT neuromuscular junction.";  
 RL Dev. Biol. 174:125-139(1996).  
 CC -1- FUNCTION: INTEGRIN ALPHA-7/BETA-1 IS THE PRIMARY LAMININ RECEPTOR  
 CC ON SKELETAL MYOBLASTS AND ADULT MYOFIBERS. DURING MYOGENIC  
 CC DIFFERENTIATION, IT MAY INDUCE CHANGES IN THE SHAPE AND MOBILITY  
 CC OF MYOBLASTS, AND FACILITATE THEIR LOCALIZATION AT LAMININ-RICH  
 CC SITES OF SECONDARY FIBER FORMATION. INVOLVED IN THE MAINTENANCE OF  
 CC THE MYOFIBERS CYTOARCHITECTURE AS WELL AS FOR THEIR ANCHORAGE,  
 CC VIABILITY AND FUNCTIONAL INTEGRITY.  
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA  
 CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAIN LINKED BY A  
 CC DISULFIDE BOND. ALPHA-7 ASSOCIATES WITH BETA-1.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS: ALPHA-7X1A, ALPHA-7X1B  
 CC (SHOWN HERE) AND ALPHA-7X1C; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN SKELETAL AND CARDIAC MUSCLE.  
 CC EXPRESSED IN REPLICATING MYOBLASTS. IN DIFFERENTIATED MUSCLE  
 CC FIBERS LOCALIZES BETWEEN FIBERS AND THE SURROUNDING MATRIX.  
 CC ISOFORMS A AND B ARE EXPRESSED AT MYOTENDINUS AND NEUROMUSCULAR  
 CC JUNCTIONS; ISOFORM C IS EXPRESSED AT NEUROMUSCULAR JUNCTIONS AND  
 CC AT EXTRASYNEPTIC SITES.  
 CC -1- DEVELOPMENTAL STAGE: ISOFORMS ARE DEVELOPMENTALLY REGULATED DURING  
 CC THE FORMATION OF SKELETAL MUSCLE. ISOFORMS A AND C ARE INDUCED  
 CC UPON TERMINAL MYOGENIC DIFFERENTIATION; ISOFORM B IS PRESENT  
 CC EARLIER IN REPLICATING CELLS AND DIMINISHES UPON DIFFERENTIATION.  
 CC -1- PTM: ADP-RIBOSYLATED ON AT LEAST TWO SITES IN THE EXTRACELLULAR  
 CC DOMAIN IN SKELETAL MYOTUBES (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.  
 CC -1- SIMILARITY: CONAINS 7 FG-GAP REPEATS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X65036; CAA46170.1;  
 DR EMBL: X74293; CAA52346.1;  
 DR EMBL: X74294; CAA52347.1;  
 DR HSP: P11215; IABX.

DR		InterPro: IPRO000413; Integrin_alpha.
DR	Pfam: PF00357; Integrin_A; 1.	
DR	Pfam: PF01839; FG-GAP; 5.	
DR	PRINTS: PR01185; INTEGRINA.	
DR	SMART: SM00191; Intl_alpha; 5.	
DR	PROSITE: PS00242; INTEGRIN_ALPHA; 1.	
KW	Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;	
KM	Repeat; Alternative splicing; ADP-ribosylation; Calcium.	
FT CHAIN	1 882	
FT FT	CHAIN 886 1106	
FT FT	DOMAIN 1 1003	
FT FT	TRANSMEM 1004 1029	
FT FT	DOMAIN 1030 1106	
FT REPEAT	17 55	
FT REPEAT	90 122	
FT REPEAT	162 194	
FT REPEAT	233 269	
FT REPEAT	290 328	
FT REPEAT	351 387	
FT REPEAT	409 449	
FT CA_BIND	300 308	
FT CA_BIND	362 370	
FT CA_BIND	420 428	
FT DOMAIN	880 885	
FT SITE	1032 1036	
FT DOMAIN	1082 1101	
FT REPEAT	1082 1085	
FT REPEAT	1090 1093	
FT REPEAT	1098 1101	
FT DISULFID	61 71	
FT DISULFID	108 131	
FT DISULFID	152 165	
FT DISULFID	467 474	
FT DISULFID	480 543	
FT DISULFID	609 615	
FT DISULFID	708 719	
FT DISULFID	866 920	
FT DISULFID	926 931	
FT CARBOHYD	713 713	
FT CARBOHYD	915 915	
FT CARBOHYD	950 950	
FT CARBOHYD	970 970	
FT VARSPLIC	1031 1106	
SO SEQUENCE	1106 AA; 121101 MW; 21B2A187837E01F6 CRC64;	
Query Match	36.9%; Score 46.5; DB 1; Length 1106;	
Best Local Similarity	47.8%; Pred. No. 32;	
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;		
OY 2 DLQERGDNDI--SPFGSGDPGF 21		
Db 362 DLNQGDFPDIAVGAFDDGKVF 384		
IITAT_MOUSE STANDARD: PRT: 1179 AA.		
ID IITAT_MOUSE AC Q61738; P70350; O88732; Q61737; O88731; Q61741;		
DT 16-OCT-2001 (Rel. 40., Created)		
DT 16-OCT-2001 (Rel. 40., Last sequence update)		
DT 15-JUN-2002 (Rel. 41, Last annotation update)		
DE Integrin alpha-7 precursor.		
GN ITGA7.		
OS Mus musculus (Mouse).		
OC Eukaryota; Metazoa; Chordata; Cranialta; Vertebrate; Euteleostomi;		
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
NCBI_taxonomy: 10090;		

RP [1] SEQUENCE FROM N.A. (ISOFORMS ALPHA-7A: ALPHA-7A2B AND ALPHA-7A1X2).  
RC STRAIN-BALB/C; TISSUE=heart;  
RA MEDLINE-94075378; PubMed-8255814;  
RX Zieber B.L., Vu M.P., Walsh N., Crawford J., Lin C.-S., Kramer R.H.;  
RT "Alternative extracellular and cytoplasmic domains of the integrin  
alpha 7 subunit are differentially expressed during development.";  
RL J. Biol. Chem. 268:26773-26783(1993).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS ALPHA-7A1X2A AND ALPHA-7A1X2B).  
RC STRAIN-129/SV;  
RA Saher G., Echtermeyer F., Beler D.R., Poeschl E., Mayer U.;  
RT "Genomic organization and chromosomal localization of the mouse  
integrin alpha7 gene.";  
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 1-70 FROM N.A.  
RC STRAIN-C57BL/6 X CBA;  
RX MEDLINE-96394366; PubMed-8798472;  
RA Zieber B.L., Kramer R.H.;  
RT "Identification and characterization of the cell type-specific and  
developmentally regulated alpha7 integrin gene promoter.";  
RL J. Biol. Chem. 271:22915-22922(1996).  
RN [4]  
RP SEQUENCE OF 34-58.  
RC TISSUE=melanoma;  
RX MEDLINE-92198982; PubMed-1839357;  
RA Kramer R.H., Vu M.P., Cheng Y.F., Ramos D.M., Timol R., Walsh N.;  
RT "Laminin-binding integrin alpha 7 beta 1: functional characterization  
and expression in normal and malignant melanocytes.";  
RL Cell Regul. 2:805-817(1991).  
RN [5]  
RP PARTIAL SEQUENCE FROM N.A. (ISOFORMS ALPHA-7A AND ALPHA-7B).  
RC STRAIN-C57BL/6 X BALB/C;  
RX MEDLINE-93366824; PubMed-8360188;  
RA Collo G., Starr L., Quaranta V.;  
RT "A new isoform of the laminin receptor integrin alpha 7 beta 1 is  
developmentally regulated in skeletal muscle.";  
RL J. Biol. Chem. 268:19019-19024(1993).  
RN [6]  
RP TISSUE SPECIFICITY.  
RX MEDLINE-96197133; PubMed-8656012;  
RA Martin P.T., Kaufman S.J., Kramer R.H., Sanes J.R.;  
RT "Synaptic integrins in developing, adult, and mutant muscle: selective  
association of alpha1, alpha7A, and alpha7B integrins with the  
neuromuscular junction."  
RL Dev. Biol. 174:125-139(1996).  
RN [7]  
RP FUNCTION.  
RX MEDLINE-98016417; PubMed-9354797;  
RA Mayer U., Saher G., Fassler R., Bornemann A., Echtermeyer F.,  
von der Mark H., Mosge N., Poeschl E., von der Mark K.;  
RT "Absence of Integrin alpha 7 causes a novel form of muscular  
dystrophy."  
RL Nat. Genet. 17:318-323(1997).  
RN [8]  
RP ADP-RYBOSYLATION.  
RX MEDLINE-95238432; PubMed-7721841;  
RA Zolkiewska A., Moss J.;  
RT "Processing of Adp-ribosylated Integrin alpha 7 in skeletal muscle  
myotubes";  
RL J. Biol. Chem. 270:9227-9233(1995).  
RN [9]  
RP -1- FUNCTION: INTEGRIN ALPHA-7/BETA-1 IS THE PRIMARY LAMININ RECEPTOR  
ON SKELETAL MYOBLASTS AND ADULT MYOTUBES. DURING MYOGENIC  
DIFFERENTIATION, IT MAY INDUCE CHANGES IN THE SHAPE AND MOBILITY  
OF MYOBLASTS, AND FACILITATE THEIR LOCALIZATION AT LAMININ-RICH  
SITES OF SECONDARY FIBER FORMATION. INVOLVED IN THE MAINTENANCE OF  
THE MYOFIBERS CYTOARCHITECTURE AS WELL AS FOR THEIR ANCHORAGE.  
CC VIABILITY AND FUNCTIONAL INTEGRITY. MICE CARRYING A TGA7 NULL  
CC ALLELE ARE VIBILE AND FERTILE. BUT SHOW PROGRESSIVE MUSCULAR  
CC DYSTROPHY STARTING SOON AFTER BIRTH. BUT WITH A DISTINCT  
CC VARIABILITY IN DIFFERENT MUSCLE TYPES.  
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA



RT subunit generated by alternative intron splicing.";  
 RL Biochem. Biophys. Res. Commun. 243:317-325(1998).  
 [2]  
 RN  
 RP SEQUENCE FROM N.A. (ISOFORM ALPHA-7X2B).  
 RX MEDLINE-98250181; PubMed-9590299;  
 RA Hayashi Y.K., Chou F.-L., Engvall E., Ogawa M., Matsuda C.,  
 RA Hirabayashi S., Yokochi K., Zieber B.L., Kramer R.H., Kaufman S.J.,  
 RA Oawa E., Goto Y.-I., Nonaka I., Tsukahara T., Wang J.Z.,  
 RA Hoffman E.P., Arakata K.;  
 RT "Mutations in the integrin alpha7 gene cause congenital myopathy.";  
 RL Nat. Genet. 19:94-97(1998).  
 [3]  
 RN  
 RP SEQUENCE FROM N.A. (ISOFORM ALPHA-7X2B).  
 RA Viziianakis I.S., Zieber B.L., Kramer R.H.;  
 RT "Cloning of human integrin alpha-7 cDNA.";  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 [4]  
 RN  
 RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.  
 RX TISSUE-Skeletal muscle; PubMed-10403775;  
 RA MEDLINE-99333684; PubMed-10403775;  
 RA Vignier N., Moghadassadeh B., Gary F., Beckmann J., Mayer U.,  
 RA Gulcener P.;  
 RT "Structure, genetic localization, and identification of the cardiac  
 RT and skeletal muscle transcripts of the human integrin alpha7 gene  
 RT (ITGA7)." ;  
 RL Biochem. Biophys. Res. Commun. 260:357-364(1999).  
 [5]  
 RN  
 RP SEQUENCE OF 34-45.  
 RX TISSUE-Melanoma; PubMed-1839357;  
 RA MEDLINE-92198982; PubMed-1839357;  
 RA Kramer R.H., Vu M.P., Cheng Y.F., Ramos D.M., Timpl R., Waleh N.;  
 RT "Laminin-binding integrin alpha 7 beta 1: functional characterization  
 RT and expression in normal and malignant melanocytes." ;  
 RL Cell Regul. 2:805-817(1991).  
 [6]  
 RN  
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORM ALPHA-7X2).  
 RX TISSUE-Heart;  
 RA MEDLINE-94075378; PubMed-8253814;  
 RA Zieber B.L., Vu M.P., Waleh N., Crawford J., Lin C.-S., Kramer R.H.;  
 RT "Alternative extracellular and cytoplasmic domains of the integrin  
 RT alpha 7 subunit are differentially expressed during development." ;  
 RL J. Biol. Chem. 268:26773-26783(1993).  
 [7]  
 RN  
 RP SEQUENCE OF 1105-1181 FROM N.A. (ISOFORM ALPHA-7B).  
 RX TISSUE-Fetal muscle;  
 RA MEDLINE-94111924; PubMed-8126096;  
 RA Song W.K., Wang W., Sato H., Bieleser D.A., Kaufman S.J.;  
 RT "Expression of alpha 7 integrin cytoplasmic domains during skeletal  
 RT muscle development: alternate forms, conformational change, and  
 RT homologues with serine/threonine kinases and tyrosine phosphatases." ;  
 RL J. Cell Sci. 106:1139-1152(1993).  
 [8]  
 RN  
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORM ALPHA-7A).  
 RX TISSUE-Skeletal muscle;  
 RA MEDLINE-98012902; PubMed-9352853;  
 RA Basora N., Vachon P.H., Herring-Gilliam F.E., Perreault N.,  
 RA Beaulieu J.-F.;  
 RT "Relation between integrin alpha7beta1 expression in human intestinal  
 RT cells and enterocytic differentiation." ;  
 RL Gastroenterology 113:1510-1521(1997).  
 [9]  
 RN  
 RP FUNCTION.  
 RX MEDLINE-97453229; PubMed-9307969;  
 RA Zieber B.L., Chen Y.Q., Kramer R.H.;  
 RT "The laminin-binding activity of the alpha 7 integrin receptor is  
 RT defined by developmentally regulated splicing in the extracellular  
 RT domain." ;  
 RL Mol. Biol. Cell 8:1723-1734(1997).  
 [10]  
 RN  
 RP FUNCTION.  
 RX MEDLINE-20160722; PubMed-10694445;  
 RA Schreiber S., Mehlitz D., Echterneyer F., Hapke S., Poeschl E.,  
 RA von der Mark H., Moch H., von der Mark K.;

RT "The role of extracellular and cytoplasmic splice domains of  
 RT alpha7-integrin in cell adhesion and migration on laminins." ;  
 RL Exp. Cell Res. 255:303-313(2000).  
 [11]  
 RN  
 RP TISSUE SPECIFICITY.  
 RX MEDLINE-96197133; PubMed-8626012;  
 RA Martin P.T., Kaufman S.J., Kramer R.H., Sanes J.R.;  
 RT "Synaptic integrins in developing, adult, and mutant muscle: selective  
 RT association of alpha1, alpha7A, and alpha7B integrins with the  
 RT neuromuscular junction." ;  
 RL Dev. Biol. 174:125-139(1996).  
 [12]  
 RN  
 RP FUNCTION: INTEGRIN ALPHA-7/BETA-1 IS THE PRIMARY LAMININ RECEPTOR  
 ON SKELETAL MYOBLASTS AND ADULT MYOFIBERS. DURING MYOGENIC  
 DIFFERENTIATION, IT MAY INDUCE CHANGES IN THE SHAPE AND MOBILITY  
 OF MYOBLASTS, AND FACILITATE THEIR LOCALIZATION AT LAMININ-RICH  
 SITES OF SECONDARY FIBER FORMATION. IT IS INVOLVED IN THE  
 MAINTENANCE OF THE MYOFIBERS CYTOARCHITECTURE AS WELL AS FOR THEIR  
 ANCHORAGE, VIABILITY AND FUNCTIONAL INTEGRITY. ISOFORMS ALPHA-7X2B  
 AND ALPHA-7X1B PROMOTE MYOBLAST MIGRATION ON LAMININ 1 AND LAMININ  
 2/4, BUT ALPHA-7X1B IS LESS ACTIVE ON LAMININ 1 (IN VITRO).  
 [13]  
 RN  
 RP SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA  
 SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAIN LINKED BY A  
 DISULFIDE BOND. ALPHA-7 ASSOCIATES WITH BETA-1.  
 [14]  
 RN  
 RP SUBCELLULAR LOCATION: Type I membrane protein.  
 [15]  
 RN  
 RP ALTERNATIVE PRODUCTS: AT LEAST 12 ISOFORMS: ALPHA-7X1A, ALPHA-  
 7X1B, ALPHA-7X1DA, ALPHA-7X1DB, ALPHA-7X2A, ALPHA-7X2B, ALPHA-  
 7X2DA, ALPHA-7X2DB, ALPHA-7X1X2A, ALPHA-7X1X2B (SHOWN HERE),  
 ALPHA-7X1X2DA AND ALPHA-7X1X2DB; MAY BE PRODUCED BY ALTERNATIVE  
 SPLICING. THERE IS A COMBINATION OF AT LEAST FIVE ALTERNATIVE  
 SPLICED DOMAINS, THREE EXTRACELLULAR (X1, X2 AND D) AND TWO  
 CYTOPLASMIC (A AND B). A THIRD POTENTIAL ALTERNATIVE SPLICED  
 CYTOPLASMIC DOMAIN (C) DOES NOT APPEAR TO BE EXPRESSED. IN VITRO  
 GENERATED ISOFORM X2C SHOWS FUNCTION. SO FAR DETECTED ARE ISOFORMS  
 ALPHA-7X1A, ALPHA-7X2B (MAJOR), ALPHA-7X2DB (MINOR) AND MINOR  
 ISOFORMS CONTAINING SEGMENT X1X2.  
 [16]  
 RN  
 RP TISSUE SPECIFICITY: ISOFORMS CONTAINING SEGMENT A ARE  
 PREDOMINANTLY EXPRESSED IN SKELETAL MUSCLE. ISOFORMS CONTAINING  
 SEGMENT B ARE ABUNDANTLY EXPRESSED IN SKELETAL MUSCLE, MODERATELY  
 IN CARDIAC MUSCLE, SMALL INTESTINE, COLON, OVARY AND PROSTATE AND  
 WEAKLY IN LUNG AND TESTES. ISOFORMS CONTAINING SEGMENT X2D ARE  
 EXPRESSED AT LOW LEVELS IN FETAL AND ADULT SKELETAL MUSCLE AND IN  
 CARDIAC MUSCLE, BUT ARE NOT DETECTED IN MYOBLASTS AND MYOTUBES. IN  
 MUSCLE FIBERS ISOFORMS CONTAINING SEGMENT A AND B ARE EXPRESSED AT  
 MOTENDINOUS AND NEUROMUSCULAR JUNCTIONS; ISOFORMS CONTAINING  
 SEGMENT C ARE EXPRESSED AT NEUROMUSCULAR JUNCTIONS AND AT  
 EXTRASYNEPTIC SITES. ISOFORMS CONTAINING SEGMENTS X1 OR X2 OR, AT  
 LOW LEVELS, X1X2 ARE EXPRESSED IN FETAL AND ADULT SKELETAL MUSCLE  
 (MYOBLASTS AND MYOTUBES) AND CARDIAC MUSCLE.  
 [17]  
 RN  
 RP DEVELOPMENTAL STAGE: IN RENOVING INTESTINAL EPITHELIUM, EXPRESSION  
 OF ISOFORMS CONTAINING SEGMENT B CORRELATES WITH THE ONSET OF  
 ENTEROCYTTIC DIFFERENTIATION.  
 [18]  
 RN  
 RP PTM: ADP-RIBOSYLATED ON AT LEAST TWO SITES OF THE EXTRACELLULAR  
 DOMAIN IN SKELETAL MYOTUBES (BY SIMILARITY).  
 [19]  
 RN  
 RP DISEASE: DEFECTS IN ITGA7 ARE ASSOCIATED WITH A FORM OF CONGENITAL  
 MYOPATHY; A GROUP OF HETEROGENEOUS MUSCLE DISORDERS WHICH ARE  
 THOUGHT TO RESULT FROM ABNORMAL MUSCLE DEVELOPMENT. MUSCLE  
 WEAKNESS IS EITHER NON-PROGRESSIVE OR SLOWLY PROGRESSIVE AND  
 APPARENT FROM BIRTH OR EARLY INFANCY.  
 [20]  
 RN  
 RP SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.  
 [21]  
 RN  
 RP SIMILARITY: CONTAINS 7 FG-GAP REPEATS.  
 [22]  
 RN  
 RP THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb.ch/announce/>  
 or send an email to [license@isb.slb.ch](mailto:license@isb.slb.ch)).  
 [23]  
 DR  
 EMBL: AF032108; AAC39708.1; -  
 DR  
 EMBL: AF032050; AAC18968.1; -  
 DR  
 EMBL: AF072132; AAC80458.1; -  
 DR  
 EMBL: AJ228936; CAB41534.1; -



```

DR EMBL: AJ228837: CAB41534.1: JOINED.
DR EMBL: AJ228838: CAB41534.1: JOINED.
DR EMBL: AJ228839: CAB41534.1: JOINED.
DR EMBL: AJ228840: CAB41534.1: JOINED.
DR EMBL: AJ228842: CAB41534.1: JOINED.
DR EMBL: AJ228843: CAB41534.1: JOINED.
DR EMBL: AJ228844: CAB41534.1: JOINED.
DR EMBL: AJ228845: CAB41534.1: JOINED.
DR EMBL: AJ228846: CAB41534.1: JOINED.
DR EMBL: AJ228847: CAB41534.1: JOINED.
DR EMBL: AJ228848: CAB41534.1: JOINED.
DR EMBL: AJ228849: CAB41534.1: JOINED.
DR EMBL: AJ228850: CAB41534.1: JOINED.
DR EMBL: AJ228851: CAB41534.1: JOINED.
DR EMBL: AJ228852: CAB41534.1: JOINED.
DR EMBL: AJ228853: CAB41534.1: JOINED.
DR EMBL: AJ228854: CAB41534.1: JOINED.
DR EMBL: AJ228855: CAB41534.1: JOINED.
DR EMBL: AJ228856: CAB41534.1: JOINED.
DR EMBL: AJ228857: CAB41534.1: JOINED.
DR EMBL: AJ228858: CAB41534.1: JOINED.
DR EMBL: AJ228859: CAB41534.1: JOINED.
DR EMBL: AJ228860: CAB41534.1: JOINED.
DR EMBL: AJ228862: CAB41534.1: JOINED.
DR EMBL: AJ228863: CAB41535.1: JOINED.
DR EMBL: AJ228867: CAB41535.1: JOINED.
DR EMBL: AJ228868: CAB41535.1: JOINED.
DR EMBL: AJ228869: CAB41535.1: JOINED.
DR EMBL: AJ228870: CAB41535.1: JOINED.
DR EMBL: AJ228871: CAB41535.1: JOINED.
DR EMBL: AJ228872: CAB41535.1: JOINED.
DR EMBL: AJ228873: CAB41535.1: JOINED.
DR EMBL: AJ228874: CAB41535.1: JOINED.
DR EMBL: AJ228875: CAB41535.1: JOINED.
DR EMBL: AJ228876: CAB41535.1: JOINED.
DR EMBL: AJ228877: CAB41535.1: JOINED.
DR EMBL: AJ228878: CAB41535.1: JOINED.
DR EMBL: AJ228879: CAB41535.1: JOINED.
DR EMBL: AJ228880: CAB41535.1: JOINED.
DR EMBL: AJ228881: CAB41535.1: JOINED.
DR EMBL: AJ228882: CAB41535.1: JOINED.
DR EMBL: AJ228883: CAB41535.1: JOINED.
DR EMBL: AJ228884: CAB41535.1: JOINED.
DR EMBL: AJ228885: CAB41535.1: JOINED.
DR EMBL: AJ228886: CAB41535.1: JOINED.
DR EMBL: AJ228887: CAB41535.1: JOINED.
DR EMBL: AJ228888: CAB41535.1: JOINED.
DR EMBL: AJ228889: CAB41535.1: JOINED.
DR EMBL: AJ228890: CAB41535.1: JOINED.
DR EMBL: AJ228891: CAB41535.1: JOINED.
DR EMBL: AJ228892: CAB41535.1: JOINED.
DR EMBL: AJ228893: CAB41535.1: JOINED.
DR EMBL: AJ228894: CAB41535.1: JOINED.
DR EMBL: AJ228895: CAB41535.1: JOINED.
DR EMBL: AJ228896: CAB41535.1: JOINED.
DR EMBL: AJ228897: CAB41535.1: JOINED.
DR EMBL: AJ228898: CAB41535.1: JOINED.
DR EMBL: AJ228899: CAB41535.1: JOINED.
DR EMBL: AJ229048: CAB41535.1: JOINED.

Query Match      36.98; Score 46.5; DB 1; Length 1181;
Best Local Similarity 47.88; Pred. No. 34;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

OY      2 DLQERGDNDI---SPFGSGDGF 21
Db      434 DLNDGFPDIAVGAHPDGDKVF 456

RESULT 7
OUEA_PSEAE
ID OUEA_PSEAE STANDARD: PRT: 347 AA.
AC 09HXH8;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE S-adenosylmethionine:RNA-ribosyltransferase-isomerase (EC 5.-.-.-)
DE (Quenosine biosynthesis protein quea).
GN OUEA OR PA3824.
OS Pseudomonas aeruginosa.
OC Bacteria: Proteobacteria: gamma subdivision: Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RC MEDLINE=20437337; PubMed=10984043;
RA Hickey C.K., Pham X.-O.T., Hwin A.L., Mizoguchi S.D., Warrenner P.,
RA Stover C.K., Brinkman F.S.L., Hufnagle W.O., Kovall D.J., Lagrou M.,
RA Garber R.L., Goltz L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.V., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RA opportunistic pathogen.";
RT Nature 406:959-964(2000).

```

```

CC CC -1- FUNCTION: Synthesizes oq from preQ1 in a single 5-
CC CC adenosylmethionine-requiring step. The ribosyl moiety of Adomet is
CC CC transferred and isomerized to the epoxycyclopentane residue of oq
CC CC (By similarity).
CC CC -1- PATHWAY: Quenosine biosynthesis.
CC CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC CC -1- SIMILARITY: BELONGS TO THE QUA FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR EMBL: AE004799: AAC07211.1: -
DR InterPro: IPR003699; Quenosine_synth.
DR Pfam: PF02347; Quenosine_synth: 1.
DR TIGRfams: TIGR00113; quea: 1.
KW Quenosine biosynthesis; Transferase; Isomerase; Complete proteome.
SQ SEQUENCE 347 AA; 38160 MW; 48ECAF074C2BC589 CRC64;

Query Match      36.18; Score 45.5; DB 1; Length 347;
Best Local Similarity 42.98; Pred. No. 12;
Matches 9; Conservative 3; Mismatches 2; Indels 7; Gaps 1;

OY      8 DNDISPSGD-----GDPF 21
Db      263 DGEIKPFGSDTDFITFGRPF 283

RESULT 8
IT4A_DROME
ID IT4A_DROME STANDARD: PRT: 1015 AA.
AC 03V7A4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Integrin alpha-PS4 precursor (Position-specific antigen 4, alpha
DE chain).
GN CG16827.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriil J.F., Agbayani A., An H.-J., Andrews-Plannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtils K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
RA de Rados B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ideyama C.,
RA Jalali B., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

```

```
QY      2 DLQERGDNDI---SPFSGDGQPF 21
          |::|::|::|::|::|
Db      343 DINKDGYNDVAVGAPFAGNGSVF 365
```

ID	CAML_MOUSE	STANDARD:	PRT:	1260 AA.
AC	CAML_MOUSE			
AD	P1627:			
DT	01-OCT-1989 (Rel. 12, Created)			
DT	01-OCT-1989 (Rel. 12, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Neural cell adhesion molecule L1 precursor (N-CAM L1).			
GN	L1CAM OR CAML1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RC	TISSUE=Brain;			
RC	MEDLINE=88318924; PubMed=34112448;			
RA	Moos M., Tacke R., Scherer H., Teplow D., Fruh K., Schachner M.;			
RT	"Neural adhesion molecule L1 as a member of the immunoglobulin superfamily with binding domains similar to fibronectin.";			
RL	Nature 334:701-703(1988).			
CC	-I- FUNCTION: CELL ADHESION MOLECULE WITH AN IMPORTANT ROLE IN THE DEVELOPMENT OF THE NERVOUS SYSTEM. INVOLVED IN NEURON-NEURON ADHESION, NERVE FASCICULATION, OUTGROWTH OF NEURITES, ETC. BINDS TO AXONIN ON NEURONS.			
CC	-I- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-I- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.			
CC	-I- SIMILARITY: CONTAINS 6 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.			
CC	-I- SIMILARITY: CONTAINS 5 FIBRONECTIN TYPE III-LIKE DOMAINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/"><u>http://www.isb-sib.ch/announce/</u></a> or send an email to <a href="mailto:license@isb-sib.ch"><u>license@isb-sib.ch</u></a> ).			
CC	-----			
DR	EMBL; X12875; CA431368.1; -.			
DR	PIR; S05479; S03479.			
DR	HSSP; P20241; 1CBP.			
DR	MGD; MGI:96721; L1cam.			
DR	InterPro; IPR003961; FN_III.			
DR	InterPro; IPR003962; FNIII_repeat.			
DR	InterPro; IPR003966; Ig_MHC.			
DR	InterPro; IPR003598; Ig_c2.			
DR	InterPro; IPR003600; Ig_like.			
DR	Pfam; PF00041; fn3; 4.			
DR	Pfam; PF00047; Ig; 6.			
DR	PRINTS; PR00014; FNTYPEIII.			
DR	SMART; SM00060; FN3; 3.			
DR	SMART; SM00410; IG_1like; 1.			
DR	SMART; SM00408; IGC2; 5.			
KW	Cell adhesion; Glycoprotein; Transmembrane; Repeat; Brain;			
KW	Immunoglobulin domain; Signal.			
FT	SIGNAL	1	19	
FT	CHAIN	20	1260	
FT	DOMAIN	20	1123	
FT	TRANSDEM	1124	1146	
FT	DOMAIN	1147	1260	
FT	DOMAIN	50	120	
FT	DOMAIN	150	215	
FT	DOMAIN	256	318	
FT	DOMAIN	346	410	
FT	DOMAIN	440	503	
FT	DOMAIN	531	599	
FT	DOMAIN	827	896	
FT	DOMAIN	932	994	
FT	DOMAIN	1032	1094	
FT	SITE	553	555	
FT	SITE	562	564	
FT	DISULFID	57	113	

```

FT DISULFID 157 208 BY SIMILARITY.
FT DISULFID 263 311 BY SIMILARITY.
FT DISULFID 353 403 BY SIMILARITY.
FT DISULFID 447 496 BY SIMILARITY.
FT DISULFID 538 590 BY SIMILARITY.
FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 478 478 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 504 504 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 587 587 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 670 670 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 725 725 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 776 776 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 824 824 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 848 848 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 875 875 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 968 968 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 978 978 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1022 1022 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1030 1030 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1073 1073 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1107 1107 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1260 AA: 140968 MW: 22857001CB2A538 CRC64;

```

```

Query Match 36.1% Score 45.5: DB 1: Length 1260;
Best Local Similarity 61.1% Pred. No. 53;
Matches 11: Conservative 2; Mismatches 4; Indels 1; Gaps 1;
OY 2 DLQERGNDISPFGDQ 19
Db 558 DLQERGNDS-KYFIEDCK 574

```

```

RESULT 10
TS2A_MOUSE STANDARD; PRT; 301 AA.
AC OGVIG3: OGDAL5:
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Testis-specific gene A2 (Male meiotic metaphase chromosome-associated
  acidic protein) (Melchroacidin).
GN TSGA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
RC TISSUE-Testis.
RX MEDLINE-98246622: Pubmed-9578619;
RA Tsuchida J., Nishina Y., Wakabayashi N., Nozaki M., Sakai Y.,
RA Nishimune Y.;
RT "Molecular cloning and characterization of melchroacidin (male meiotic
  metaphase chromosome-associated acidic protein).";
RL Dev. Biol. 197:67-76(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6J: TISSUE-Testis;
RX MEDLINE-21085660: Pubmed-11217851;
RA Atakawa T., Hara A., Shihata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishikuni Y., Konno H., Adachi J., Fukuda S.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,

```

```

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guinacino S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Marzelli J., Mombert P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Suzuki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Hayashizaki Y.;
RA Wyszynski A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- FUNCTION: The specific expression during male germ cell
CC development and its characteristic localization suggest that it
CC may play an important role in male meiosis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic in late spermatocytes, secondary
CC spermatocytes and round spermatids. Gathered around metaphase
CC chromosomes during meiotic divisions.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be
CC produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: Germ cell specific. Specifically expressed in
CC testis, and to a lower extent in ovary. Not expressed in somatic
CC tissues.
CC -1- DEVELOPMENTAL STAGE: During male germ cell development it is not
CC detected until 12 days. Significant expression is first detected
CC 14-day-old through to adult testis. Expression is first detected
CC in the pachytene spermatocytes at stage V, becomes stronger from
CC the late pachytene spermatocytes to round spermatid stage, and
CC then gradually decreases as the morphogenesis proceeds further.
CC Not expressed in germ cells located in the first layer of the
CC seminiferous epithelium (spermatogonia, leptotene and zygotene
CC spermatocytes).
CC -1- SIMILARITY: CONTAINS 6 MORN REPEATS.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 283.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB005635; BAB83693.1; ALT_FRAME.
DR MGD: AK005739; BAB24214.1; -.
DR MGD: MGI:1194909; Tsga2.
DR InterPro: IPR003409; MORN.
DR Pfam: PF02493; MORN; 6.
DR KEGG: Repeat; Repeat; Alternative splicing.
FT DOMAIN 195 201 POLY-GLU
FT REPEAT 20 43 MORN REPEAT 1.
FT REPEAT 44 66 MORN REPEAT 2.
FT REPEAT 67 89 MORN REPEAT 3.
FT REPEAT 90 112 MORN REPEAT 4.
FT REPEAT 113 135 MORN REPEAT 5.
FT REPEAT 159 181 MORN REPEAT 6.
FT REPEAT 181 181 MORN REPEAT 6.
FT REPEAT 181 181 MORN REPEAT 6.
SQ SEQUENCE 301 AA: 34181 MW: 704ABEF994DFC90B CRC64;

```

```

Query Match 35.7% Score 45: DB 1: Length 301;
Best Local Similarity 37.5% Pred. No. 12;
Matches 6: Conservative 6; Mismatches 4; Indels 0; Gaps 0;
OY 2 DLQERGNDISPFGD 17
Db 8 ELKEEGNDLCEYGE 23

```

```

RESULT 11
TS2A_HUMAN STANDARD; PRT; 309 AA.
AC OGVIR4:
DT 15-JUN-2002 (Rel. 41, Created)

```

DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Testis-specific gene A2 (Male meiotic metaphase chromosome-associated  
 DE acidic protein) (Melchrocidin).  
 GN TSGA2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Testis;  
 RX MEDLINE=98246622; PubMed=9578619;  
 RA Tsuchida J., Nishina Y., Wakabayashi N., Nozaki M., Sakai Y.,  
 RA Nishimune Y.;  
 RT Molecular cloning and characterization of melchrocidin (male meiotic  
 RT metaphase chromosome-associated acidic protein)."  
 RL Dev. Biol. 197;67-76(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Testis;  
 RA Shimizu N., Kudoh J., Shibuya K.;  
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: May play an important role in male meiosis (By  
 CC similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. In late spermatocytes, secondary  
 CC chromosomes during meiotic divisions. Gathered around metaphase  
 CC -1- TISSUE SPECIFICITY: Testis specific.  
 CC -1- SIMILARITY: CONTAINS 6 MORN REPEATS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AB006536; BAB83694.1; -;  
 CC EMBL: AB041016; BAB92995.1; -;  
 CC Genev; HGNC:12371; TSGA2.  
 CC InterPro: IPR003409; MORN.  
 CC Pfam: PF02493; MORN; 6.  
 CC  
 CC KW Melosis; Repeat.  
 CC FT DOMAIN 195 201 POLY-GLU.  
 CC FT DOMAIN 228 301 POLY-GLU.  
 CC FT REPEAT 20 43 MORN REPEAT 1.  
 CC FT REPEAT 44 66 MORN REPEAT 2.  
 CC FT REPEAT 67 89 MORN REPEAT 3.  
 CC FT REPEAT 90 112 MORN REPEAT 4.  
 CC FT REPEAT 113 135 MORN REPEAT 5.  
 CC FT REPEAT 159 181 MORN REPEAT 6.  
 CC SEQUENCE 309 AA; 35124 MW; 09A20A76D5A6AB8 CRC64;  
 SQ  
 Query Match 35.7%; Score 45; DB 1; Length 309;  
 Best Local Similarity 46.7%; Pred. No. 12;  
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
 QY 2 DLQERDNDISPSG 16  
 DB 8 ELEEGENDIGEYEG 22  
 DB  
 RESULT 12  
 GUN1\_USTMA STANDARD; PRT; 393 AA.  
 ID GUN1\_USTMA  
 AC P54424;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Endoglucanase I precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase 1)  
 DE (Cellulase 1) (EG 1).

GN EGLI.  
 OS Ustilago maydis (Smut fungus).  
 OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;  
 OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.  
 OX NCBI\_TaxID=5270;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SRRAIN-PBD11;  
 RX MEDLINE=96145728; PubMed=8590631;  
 RA Schaevecker F., Manner G., Kahmann R.;  
 RT Filament-specific expression of a cellulase gene in the dimorphic  
 RT fungus Ustilago maydis."  
 RL Biol. Chem. Hoppe-Seyler 376:617-625(1995).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic  
 CC linkages in cellulose.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: HYPHAL TIP.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN FILAMENTOUS DIKARYON.  
 CC -1- PTM: MAY ALSO BE O-GLYCOSYLATED.  
 CC -1- SIMILARITY: BELONGS TO CELLULASE FAMILY K (FAMILY 45 OF GLYCOSYL  
 CC HYDROLASES).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: S81598; AAB36147.1; -;  
 CC HSSP: P43316; 2ENG.  
 CC InterPro: IPR000334; GH\_45.  
 CC Pfam: PF02015; GLYCOSYL\_HYDROL\_45; 1.  
 CC PROSITE: PS01140; GLYCOSYL\_HYDROL\_F45; 1.  
 CC KW Cellulose degradation; Hydrolase; Glycosidase; Signal; Glycoprotein.  
 CC FT SIGMUL 1 26 POTENTIAL.  
 CC FT CHAIN 27 393 ENDOLUCANASE 1.  
 CC FT ACT\_SITE 34 34 NOCLEOPHILE (BY SIMILARITY).  
 CC FT ACT\_SITE 152 152 PROTON DONOR (BY SIMILARITY).  
 CC FT DOMAIN 270 385 ALA/GLY/SER-RICH.  
 CC FT CARBOHYD 343 343 N-LINKED (GLCNAC... (POTENTIAL).  
 CC SEQUENCE 393 AA; 39594 MW; 65C753C610CD6AD3 CRC64;  
 SQ  
 Query Match 35.7%; Score 45; DB 1; Length 393;  
 Best Local Similarity 61.5%; Pred. No. 16;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 4 QERGDNDISPSG 16  
 DB 235 QQRDNDTISPSG 247  
 DB  
 RESULT 13  
 SABC\_BACAM  
 ID SABC\_BACAM STANDARD; PRT; 472 AA.  
 AC P21130;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE levansucrase precursor (EC 2.4.1.10) (Beta-D-fructofuranosyl  
 DE transferase) (Sucrose 6-fructosyl transferase).  
 GN SABC.  
 OS Bacillus amyloliquefaciens.  
 OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1390;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 23844;  
 RX MEDLINE=91092506; PubMed=2265762;  
 RA Tang L.B., Lenstra R., Borchert T.V., Vasantha N.;  
 RT Isolation and characterization of levansucrase-encoding gene from  
 RT Bacillus amyloliquefaciens."

RL Gene 96:89-93(1990).  
 CC -1- CATALYTIC ACTIVITY: Sucrose + ((2,6)-beta-D-fructosyl)(N) -  
 CC glucose + ((2,6)-beta-D-fructosyl)(N+1).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- INDUCTION: BY SUCROSE.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 68 OF GLYCOSYL HYDROLASES.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL, X52988; CAA37179.1; -  
 CC PIR, J00802; J00802.  
 CC PIR, S11739; S11739.  
 CC InterPro: IPR003469; Glyco\_hydro\_68.  
 CC Pfam: PF02435; Glyco\_hydro\_68; 1.  
 CC Transferase: Glycosyltransferase; Signal.  
 CC SIGNAL  
 CC FT CHAIN 1 29  
 CC SO SEQUENCE 472 AA; 52859 MW; F38592D27267E7D CRC64;  
 Query Match 35.7%; Score 45; DB 1; Length 472;  
 Best Local Similarity 45.0%; Pred. No. 20;  
 Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;  
 Oy 4 QERGNDISPGDCGPFKD 23  
 Db 126 QKVGNDISDMKNGRVED 145  
 ID VLL\_HPV60 STANDARD; PRT; 508 AA.  
 AC P50821; Q80947;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Major capsid protein L1.  
 GN L1.  
 OS Human papillomavirus type 60.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 OC Papillomavirus.  
 OC NCBI\_TaxID=40540;  
 RN NCBI\_TaxID=40540;  
 RA SEQUENCE FROM N.A.  
 RA Delius H.;  
 RN Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 372-467 FROM N.A.  
 RA Chan S.Y., Delius H., Halpern A.L., Bernard H.U.;  
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL, U31792; AAA9491.1; -  
 CC EMBL, U21876; AAA92837.1; -  
 CC InterPro: IPR002210; PV\_capsid\_L1.  
 CC Pfam: PF00500; late\_protein\_L1; 1.  
 CC PRINTS: PR00865; HPVcapsidL1.  
 CC ProDom: PD000544; PV\_capsid\_L1; 1.  
 CC Coat protein: Late protein.  
 CC SEQUENCE 508 AA; 57827 MW; 314954985534D00 CRC64;  
 FT

Query Match 35.7%; Score 45; DB 1; Length 508;  
 Best Local Similarity 64.3%; Pred. No. 22;  
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 Oy 2 DLQERNDISPS 15  
 Db 459 DLQERFNSLSQFS 472  
 ID YDYG\_SCHPO STANDARD; PRT; 851 AA.  
 AC Q10366;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein C22E12.16c in chromosome I.  
 GN SPAC22E12.16c.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 OC NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=972;  
 RC MEDLINE=21848401; PubMed=11859360;  
 RX Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.C., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grynoprez B.,  
 RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Flicke C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McComble W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 FT "The genome sequence of Schizosaccharomyces pombe."  
 RL Nature 415:871-880(2002).  
 CC  
 CC -1- SIMILARITY: STRONG, TO YEAST PIK1.  
 CC  
 CC -1- SIMILARITY: BELONGS TO THE PI3/P14-KINASES FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL, Z70043; CAA93903.1; -  
 CC InterPro: IPR00403; PI3\_P14\_kinase.  
 CC Pfam: PF00454; PI3\_P14\_kinase; 1.  
 CC SMART: SM00146; PI3K; 1.  
 CC PROSITE: PS00915; PI3\_4\_KINASE\_1; 1.  
 CC PROSITE: PS00916; PI3\_4\_KINASE\_2; 1.  
 CC PROSITE: PS0290; PI3\_4\_KINASE\_3; 1.  
 CC Hypothetical protein: Transferase; Kinase.  
 CC DOMAIN 581 827  
 FT PI3K/PIAK.

SQ SEQUENCE 851 AA; 96657 MM; A991F3C7E3D980BE CRC64;

Query Match 35.7%; Score 45; DB 1; Length 851;

Best Local Similarity 50.0%; Pred. No. 40;

Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 TDLQERGDNDISPFGDG 18

Db 382 TNLQDSTDNDISESESESG 399

Search completed: January 17, 2003, 13:18:02  
Job time : 11 secs

Gencore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: January 17, 2003, 13:16:38 ; Search time 29 Seconds

(without alignments)  
163.417 Million cell updates/sec

Title: US-09-641-034-47

Sequence: 126  
1 TLQERGDNDISFGDGPFD 23

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: SP archaea: \*  
2: SP bacteria: \*  
3: SP fungi: \*  
4: SP human: \*  
5: SP invertebrate: \*  
6: SP mammal: \*  
7: SP mhc: \*  
8: SP organelle: \*  
9: SP phage: \*  
10: SP plant: \*  
11: SP rodent: \*  
12: SP virus: \*  
13: SP vertebrate: \*  
14: SP unclassified: \*  
15: SP viirus: \*  
16: SP bacteriap: \*  
17: SP archaeap: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	126	100.0	525	4	09N076
2	123	97.6	500	6	095K65
3	123	97.6	555	6	09GMI3
4	123	97.6	555	6	09N076
5	81	64.3	433	11	092411
6	81	64.3	433	11	09ES02
7	52	41.3	320	2	052676
8	49.5	39.9	793	5	09VYV0
9	49.5	38.9	752	4	09C091
10	48	38.1	277	16	08Y183
11	48	38.1	385	12	092432
12	48	38.1	440	2	09F5V2
13	48	38.1	484	2	044497
14	48	38.1	484	2	09NK94
15	47	37.3	444	16	09K8J7
16	47	37.3	477	2	08RSS1

## ALIGNMENTS

17	47	37.3	481	16	08ZBM6	08ZBM6 yersinia pe
18	47	37.3	726	16	08XR29	08XR29 ralatonia s
19	47	37.3	825	16	0997N0	0997N0 staphylococ
20	46	36.5	279	2	055321	055321 synechococ
21	46	36.5	365	16	08U6F1	08U6F1 agrobacteri
22	45.5	36.1	265	2	033643	033643 salmonella
23	45.5	36.1	265	2	047220	047220 escherichia
24	45.5	36.1	1259	11	090Y38	090Y38 mus muscu
25	45	35.7	158	16	08Y4V1	08Y4V1 listeria mo
26	45	35.7	180	5	09NDR9	09NDR9 branchiosto
27	45	35.7	207	11	08Y1K4	08Y1K4 mus muscu
28	45	35.7	237	11	0912W8	0912W8 mus muscu
29	45	35.7	357	10	09S1M8	09S1M8 arabidopsis
30	45	35.7	536	12	08VA07	08VA07 chikungunya
31	45	35.7	690	4	096MB5	096MB5 homo sapien
32	45	35.7	913	16	09P003	09P003 ureaplasma
33	45	35.7	1052	2	09MHL5	09MHL5 methylolact
34	45	35.7	1419	5	09VIL3	09VIL3 drosophila
35	44.5	35.3	563	10	004367	004367 sambucus n1
36	44.5	35.3	698	16	08X4L8	08X4L8 escherichia
37	44	34.9	139	16	0983F0	0983F0 rhizobium l
38	44	34.9	156	16	08XUS9	08XUS9 ralatonia s
39	44	34.9	215	5	0819T8	0819T8 aedes aegypt
40	44	34.9	216	3	08X041	08X041 neurospora
41	44	34.9	475	16	08Z9B0	08Z9B0 salmonella
42	44	34.9	533	5	094677	094677 plasmodium
43	44	34.9	572	5	09GSR0	09GSR0 plasmodium
44	44	34.9	682	2	09ACJ4	09ACJ4 pseudomonas
45	44	34.9	683	16	0914P3	0914P3 pseudomonas

## RESULT 1

09N076 PRELIMINARY; PRT; 525 AA.  
ID 09N076;  
AC 09N076;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 15, Last annotation update)  
DE Matrix extracellular phosphoglycoprotein precursor.  
GN MEPE.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BONE.  
RA MEDLINE-20399567; PubMed-10945470;  
RA Rowe P.S.N., De Zoysa P.A., Dong R., Wang H.R., White K.E.,  
RA Econs M.J., Oudet C.L.;  
RT \*MEPE, a new gene expressed in bone marrow and tumors causing  
RT osteomalacia.\*;  
RL Genomics 67:54-68(2000) ✓  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MEDLINE-21309068; PubMed-11414762;  
RA Argiro L., Desbarats M., Glorieux F.H., Ecartot B.;  
RT \*Mepe, the gene encoding a tumor-secreted protein in oncogenic  
RT hypophosphatemic osteomalacia, is expressed in bone.\*;  
RL Genomics 74:342-351(2001).  
DR EMBL: AJ276396; CAB97250.1; -;  
DR EMBL: AF325916; AAK70343.1; -;  
KW Signal.  
FT SIGNAL.  
FT CHAIN 18 525  
SQ SEQUENCE 525 AA: 58419 MW: 0977CA6E871CA9E5 CRC64:  
POTENTIAL.  
MATRIX EXTRACELLULAR PHOSPHOGLYCOPROTEIN.  
Query Match 100.0%; Score 126; DB 4; Length 525;  
Best Local Similarity 100.0%; Pred. No. 4.4e-11;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDLQERGDNDISPFSGDQPFKD 23  
 DB 242 TDLQERGDNDISPFSGDQPFKD 264

## RESULT 2

O95KG5 PRELIMINARY: PRT; 500 AA.  
 AC 095KG5:  
 DT 01-DEC-2001 (TRENBLREL. 19, Created)  
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)  
 DE Hypothetical 55.6 kDa protein.  
 OS Macaca fascicularis (Grab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 RN NCBI\_TaxID=9541;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=TEMPORAL LOBE RIGHT;  
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,  
 RA Suzuki Y., Sugano S., Hashimoto K.;  
 RT \*Isolation of full-length cDNA clones from macaque brain cDNA  
 RT libraries.\*;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB060891; BAB46894.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 500 AA; 55577 MW; 918D265AD8EDC7BC CRC64;

Query Match 97.6%; Score 123; DB 6; Length 500;  
 Best Local Similarity 95.7%; Pred. No. 1.2e-10;  
 Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDLQERGDNDISPFSGDQPFKD 23  
 DB 218 TDLQERGDNDISPFSGDQPFKD 240

## RESULT 3

O9GM13 PRELIMINARY: PRT; 555 AA.  
 AC 09GM13:  
 DT 01-MAR-2001 (TRENBLREL. 16, Created)  
 DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)  
 DE MEPE protein.  
 OS Macaca fascicularis (Grab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 RN NCBI\_TaxID=9541;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRAIN PARIENTAL LOBE;  
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,  
 RA Suzuki Y., Sugano S., Hashimoto K.;  
 RT \*Isolation of full-length cDNA clones from macaque brain cDNA  
 RT libraries.\*;  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB050259; BAB17010.1; -.  
 SQ SEQUENCE 555 AA; 61950 MW; BD4M47E88038A9E2 CRC64;

Query Match 97.6%; Score 123; DB 6; Length 555;  
 Best Local Similarity 95.7%; Pred. No. 1.4e-10;  
 Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDLQERGDNDISPFSGDQPFKD 23  
 DB 273 TDLQERGDNDISPFSGDQPFKD 295

RESULT 4  
 O9N076 PRELIMINARY: PRT; 555 AA.  
 AC 09N076:  
 DT 01-OCT-2000 (TRENBLREL. 15, Created)  
 DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)  
 DE Matrix extracellular phosphoglycoprotein.  
 OS Macaca fascicularis (Grab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 RN NCBI\_TaxID=9541;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=TEMPORAL LOBE RIGHT;  
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,  
 RA Suzuki Y., Sugano S., Hashimoto K.;  
 RT \*Isolation of full-length cDNA clones from macaque brain cDNA  
 RT libraries.\*;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB046056; BAB01638.1; -.  
 KW Matrix extracellular phosphoglycoprotein precursor.  
 SQ SEQUENCE 555 AA; 61979 MW; CCFE1A98ADA19EB4 CRC64;

Query Match 97.6%; Score 123; DB 6; Length 555;  
 Best Local Similarity 95.7%; Pred. No. 1.4e-10;  
 Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDLQERGDNDISPFSGDQPFKD 23  
 DB 273 TDLQERGDNDISPFSGDQPFKD 295

RESULT 5  
 O92411 PRELIMINARY: PRT; 433 AA.  
 AC 092411:  
 DT 01-DEC-2001 (TRENBLREL. 19, Created)  
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)  
 DE Matrix extracellular phosphoglycoprotein precursor.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 RN NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J;  
 RX MEDLINE=21309068; PubMed=11414762;  
 RA Argiro L., Desparats M., Glorieux F.H., Ecarot B.;  
 RT \*Mepe, the gene encoding a tumor-secreted protein in oncogenic  
 RT hypophosphatemic osteomalacia, is expressed in bone.\*;  
 RL Genomics 74:342-351 (2001).  
 DR EMBL: AF314964; AAK70342.1; -.  
 DR MGD: MGI:2137384; Mepe.  
 KW Signal.  
 FT SIGNAL. 1 18 POTENTIAL.  
 FT CHAIN 19 433 MATRIX EXTRACELLULAR PHOSPHOGLYCOPROTEIN.  
 SQ SEQUENCE 433 AA; 45984 MW; 7CD603CAAB8BA1B0 CRC64;

Query Match 64.3%; Score 81; DB 11; Length 433;  
 Best Local Similarity 75.0%; Pred. No. 0.00022;  
 Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 TDLQERGDNDISPFSGDQPFKD 21  
 DB 171 TDLQERGDNDISPFSGDQPFKD 190

RESULT 6  
 O9ES02 PRELIMINARY: PRT; 435 AA.  
 AC 09ES02:



```

DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Osteoregulin.
OS Rattus norvegicus (Rat.).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=20549633; PubMed=10967096;
RA Petersen D.N., Tkalcic G.T., Mansolf A.L., Rivera-Gonzalez R.,
RA Brown T.A.,
RT "Identification of Osteoblast/Osteocyte Factor 45 (OF45), a Bone-
RT Specific cDNA Encoding an RGD-containing Protein That Is Highly
RT Expressed in Osteoblasts and Osteocytes."
RL J. Biol. Chem. 275:36172-36180(2000).
DR EMBL: AF260922; AAC33366.1; -
SQ SEQUENCE 435 AA; 46515 MW; D587F82968a26BCB CRC64;

Query Match 64.3%; Score 81; DB 11; Length 435;
Best Local Similarity 75.0%; Pred. No. 0.00037;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 DLQERGDNDISPFSGDCOPF 21
DB 165 DLVGRGDNDVPEFSGDCQHF 184

RESULT 7
O52676 PRELIMINARY; PRT: 320 AA.
AC 032676;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE Porin.
GN PORC.
OS Rhodobacter capsulatus (Rhodospseudomonas capsulata).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Rhodobacter.
OX NCBI_TaxID=1061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=37B4;
RX MEDLINE=97149280; PubMed=8996088;
RA Trileschmann M.D., Pattus F., Tadros M.H.;
RT "Molecular characterization and organization of porin from Rhodobacter
RT capsulatus strain 37B4."
RL Gene 183:61-68(1996).
DR EMBL: U57653; AAB41301.1; -
DR HSSP: P31243; 2POR.
SQ SEQUENCE 320 AA; 33396 MW; B0E1745CB9D86F93 CRC64;

Query Match 41.3%; Score 52; DB 2; Length 320;
Best Local Similarity 58.8%; Pred. No. 7;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 TDLQERGDNDISPFSGD 17
DB 112 TDLQERGDNDISPFSGD 128

RESULT 8.
O9VYV0 PRELIMINARY; PRT: 793 AA.
AC 09VYV0;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE CG14085 protein.
GN CG14085.

```

```

OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Ibbegan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclib J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Sliker E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Stryker R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003517; AAF49175.1; -
DR FlyBase: FBgn0036859; CG14085.
SQ SEQUENCE 793 AA; 88674 MW; FFC0E3DF3E537C41 CRC64;

Query Match 39.3%; Score 49.5; DB 5; Length 793;
Best Local Similarity 61.1%; Pred. No. 49;
Matches 11; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

OY 7 GDNDISPFSG-DCOPFKD 23
DB 693 GDNDISPFSGICQCNKE 710

RESULT 9
O9C091 PRELIMINARY; PRT: 752 AA.
AC 09C091;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE KIAA1772 protein (Fragment).
GN KIAA1772.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;

```

(1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-21082932; PubMed-11214970;  
 RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XIX.  
 The complete sequences of 100 new cDNA clones from brain which code  
 for large proteins in vitro.";  
 RL DNA Res. 7:347-355(2000).  
 DR EMBL: AB051559; BAB21863.1;  
 DR InterPro: IPR001969; Aspartase\_site.  
 DR PROSITE: PS00141; ASP\_PROTEASE; UNKNOWN\_1.  
 FT NON\_TER  
 SQ SEQUENCE 752 AA; 82489 MW; 7BEA604AA833D2E4 CRC64;

Query Match 38.9%; Score 49; DB 4; Length 752;  
 Best Local Similarity 57.1%; Pred. No. 55;  
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 4 QERGNDISPFGD 17  
 |||: |||  
 DB 630 QRRGDSVTFPDGD 643

## RESULT 10

OBY183 PRELIMINARY; PRT: 277 AA.  
 AC QBY183;  
 DT 01-MAR-2002 (Tremblrel. 20, Created)  
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
 DE Lysosyme M1 precursor (EC 3.2.1.17).  
 GN BME10562.  
 OS Brucella melitensis.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Brucellaceae; Brucella.  
 OX NCBI\_TaxID-29459;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-16M / ATCC 23456 / BIOTYPE 1;  
 RX MEDLINE-20020109; PubMed-11756688;  
 RA DelVecchio V.G., Kapatal V., Redkar R.J., Patra G., Mujer C., Los T.,  
 RA Ivanova N., Anderson I., Bhattacharya A., Lykdis A., Resnik G.,  
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,  
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Jettsson J.-J.,  
 RA Haselkorn R., Kyrides N., Overbeek R.;  
 RT "The genome sequence of the facultative intracellular pathogen  
 Brucella melitensis.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).  
 DR EMBL: AE009498; AL51743.1;  
 DR InterPro: IPR002053; CH\_25.  
 DR Pfam: PF01183; Glyco\_hydro\_25; 1.  
 DR ProDom: PD004620; GH\_25; 1.  
 DR Hydrolase; Glycosidase; Complete proteome.  
 SQ SEQUENCE 277 AA; 31766 MW; B8ICF4BA63E2B80 CRC64;

Query Match 38.1%; Score 48; DB 16; Length 277;  
 Best Local Similarity 52.9%; Pred. No. 25;  
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 6 RGNDISPFGDQPFK 22  
 |||: |||  
 DB 252 RGDADINTFAGDSASWK 268

## RESULT 11

O92432 PRELIMINARY; PRT: 385 AA.  
 AC O92432;  
 DT 01-NOV-1998 (Tremblrel. 08, Created)  
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE LEF-3-ACMNPV orf67.  
 GN LEF-3.

OS Bombyx mori nuclear polyhedrosis virus (BmNPV).  
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
 OC Nucleopolyhedrovirus.  
 OX NCBI\_TaxID-10458;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-T3;  
 RA MEDLINE-97329351; PubMed-9185864;  
 RA Kamita S.G., Maeda S.;  
 RT "Sequencing of the putative DNA helicase-encoding gene of the Bombyx  
 mori nuclear polyhedrosis virus and fine-mapping of a region involved  
 in host range expansion.";  
 RL Gene 190:173-179(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-T3;  
 RX MEDLINE-99281911; PubMed-10355780;  
 RA Gomi S., Majima K., Maeda S.;  
 RT "Sequence analysis of the genome of Bombyx mori  
 nucleopolyhedrovirus.";  
 RL J. Gen. Virol. 80:1323-1337(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-T3;  
 RA Maeda S.;  
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-T3;  
 RA Gomi S.;  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: L33180; AAC63740.1;  
 SQ SEQUENCE 385 AA; 44875 MW; 1B2F1A12EB8207BB CRC64;

Query Match 38.1%; Score 48; DB 12; Length 385;  
 Best Local Similarity 56.2%; Pred. No. 36;  
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 8 DNDISPFGDQPFKD 23  
 |||: |||  
 DB 61 DNKQEVYGGDSQSFKD 76

## RESULT 12

O9F5V2 PRELIMINARY; PRT: 440 AA.  
 AC O9F5V2;  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE RteB.  
 GN RteB.  
 OS Bacteroides fragilis.  
 OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Bacteroidaceae;  
 OC Bacteroides.  
 OX NCBI\_TaxID-817;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TRANSPOSON-CTN25285;  
 RA Smith C.J., Bayley D.;  
 RT "The putative IS1224 insertion sequence is actually part of a large  
 genetic element with similarity to the Bacteroides conjugative  
 transposons.";  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS A SIGMA-54 FACTOR INTERACTION ATP-BINDING  
 DOMAIN.  
 DR EMBL: AF303552; AAG17462.1;  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR002197; HTH\_Fis.  
 DR InterPro: IPR001789; Response\_reg.  
 DR InterPro: IPR002078; Sig54\_interact.  
 DR Pfam: PF02954; HTH\_8; 1.  
 DR Pfam: PF00072; response\_reg; 1.



DR PROSITE; PS00137; SUBTILASE\_HIS; 1.  
 DR PROSITE; PS00138; SUBTILASE\_SER; 1.  
 KW Protease; Complete proteome.  
 SQ SEQUENCE 444 AA; 48916 MW; 3805F81C53A21F2F CRC64;

Query Match 37.3%; Score 47; DB 16; Length 444;  
 Best Local Similarity 47.1%; Pred. No. 61;  
 Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 2 DLQERGDNDISPRSGDG 18  
 I : I : I : I : I : I : I : I :  
 Db 319 DTTDREDDVAVPFSSRG 335

Search completed: January 17, 2003, 13:18:38  
 Job time : 30 secs